2009

Cross-validation in model-assisted estimation

Lifeng You
Iowa State University

Cross-validation in model-assisted estimation

by

Lifeng You

A dissertation submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Major: Statistics

Program of Study Committee:
Jean D. Opsomer, Major Professor
Michael Larsen
Huaiqing Wu
Cindy L. Yu
Helen H. Jensen

Iowa State University

Ames, Iowa

2009

DEDICATION

I would like to dedicate this thesis to my husband Lei without whose support I would not have been able to complete this work. I would also like to thank my friends and family for their loving guidance and financial assistance during the writing of this work.

TABLE OF CONTENTS

LIST OF TABLES

ACKNOWLEDGEMENTS

I would like to take this opportunity to express my thanks to my major professor, Dr. Opsomer, for his guidance, patience and support throughout this research and the writing of this thesis. I would also like to thank my committee members for their efforts and contributions to this work: Dr. Larsen, Dr. Wu, Dr. Yu and Dr. Jensen. I would additionally like to thank Dr. Hofmann for her guidance throughout the initial stages of my graduate study. Finally I would like to thank Dr. Kling and Dr. Maiti for their valuable comments and suggestions to my research study.

ABSTRACT

Variance estimation for survey estimators that include modeling relies on approximations that ignore the effect of fitting the models. Cross-validation (CV) criterion provides a way to incorporate this effect. We will show 4 ways in which we explore this in this dissertation.

Penalized spline regression, as a main type of nonparametric model assisted methods, is a common technique to improve the precision of finite population estimators. In Chapter 1, we propose a CV based criterion to select the smoothing parameter for the penalized spline regression estimator. The design-based asymptotic properties of the method are derived, and simulation studies show how well it works in practice.

Regression estimator is a common technique to improve the precision of finite population estimators by using the available auxiliary information of the population. In Chapter 2, we propose a CV based variance estimator and compare it to other two variance estimators. The design-based asymptotic properties of the estimator are derived, and simulation studies show how well it works in practice.

Regression estimator works well for the cases where there is a strong linear relationship between regressor and regressands. On the contrary, when the relationship is weak, π estimator is a good choice. In Chapter 3, a new estimator as a linear combination of those two estimators is proposed to select between them. We introduce a CV based variance estimator for the new proposed estimator. The design-based asymptotic properties of the estimator are explored, and simulation studies show how well it works in practice.

In linear regression estimation, how to choose the set of control variables $\boldsymbol{x}$ is a difficult practical problem. In Chapter 4, a CV criterion is introduced for choosing between combinations of the $\boldsymbol{x}$ variables to be included in the model. The design-based asymptotic properties of the estimator are explored, and simulation studies show how well it works in practice.

CHAPTER 1. Cross-Validation in Penalized Spline Model-Assisted Estimation

1.1 Introduction

In many surveys, the available auxiliary information for the population can be used to improve the precision of design-based estimators. Thereinto ratio and regression estimators have been used for a long time in survey estimation, e.g., Cochran (1977).

Breidt and Opsomer (2000) proposed a nonparametric model-assisted regression estimator with the relationship between the variables to be any smooth function. They used kernel-based local polynomial regression and showed that the nonparametric estimator has the same asymptotic design properties of the parametric model-assisted estimators.

The practical properties of the estimator depend on the choice of a smoothness tuning parameter, i.e., the bandwidth in local polynomial regression. In Breidt and Opsomer (2000), the bandwidth is treated as a fixed quantity and the issue of how to best select a bandwidth value is not addressed. In Opsomer and Miller (2005), the issue of smoothing parameter selection for nonparametric model-assisted estimation was explored and a sample-based criterion that could be used for this purpose was proposed. The proposed smoothing parameter selection method was based on minimizing a type of cross-validation criterion, suitably adjusted for the effect of the finite population setting and the survey design.

Penalized spline regression, often called P-splines, is a main type of nonparametric model-assisted methods introduced by Eilers and Marx (1996). P-splines are flexible and can be incorporated into a wide range of modelling contexts. Ruppert et al. (2003) gave an overview of applications of P-splines to different settings. P-splines are also a natural candidate for constructing nonparametric small area estimators in terms of their close connections with linear mixed models discussed in Wand (2003).

The ability of combing nonparametric regression and mixed model regression with P-splines was used in different contexts, e.g., Parise et al. (2001) and Coull et al. (2001). They all provided examples of using penalized splines in the construction of mixed effect regression models to analyze the data with random effects. In the survey context, Zheng and Little (2003) proposed a model-based estimator for cluster sampling, where the regression model combines a spline model with a random effect for the clusters. Opsomer et al. (2008) proposed a new small area estimation approach, which combines small area random effects with a smooth nonparametrically specified trend. The small area estimation problem could be expressed as a mixed effect model regression by using penalized splines as the representation for the nonparametric trend. They showed consistency of the estimator, computed its mean squared error and provided tests for small area effects and non-linearities.

Breidt et al. (2005) proposed a class of estimators based on penalized spline regression. Those estimators are weighted linear combinations of sample observations, and weights are calibrated to known control totals. The estimators are design consistent and asymptotically normal under the conditions of standard design, and they admit consistent variance estimation by using design-based methods. Breidt et al. (2005) considered data-driven penalty selection in the context of unequal probability sampling designs and showed that the estimators are more efficient than parametric regression estimators when the parametric model is incorrectly specified, while being approximately same efficient when the parametric model specification is correct.

Modelling a regression function as a piecewise polynomial with a large number of pieces relative to the sample size is involved in regression spline smoothing. Since the number of possible models is so large, efficient strategies are required for choosing among them. Wand (2000) reviewed some approaches to this problem and compared them through a simulation study. For simplicity, Wand (2000) considered the univariate smoothing setting with Gaussian noise and the truncated polynomial regression spline basis. Several other approaches for knot selection exist, e.g., the TURBO algorithm in Friedman and Silverman (1989) and its subsequent generalization, the MARS algorithm in Friedman (1991).

The knots of a penalized spline are generally at fixed quantiles of the independent variable, and the only parameters that can be chosen to adjust are the number of knots and the penalty parameter. Ruppert (2002) studied the effects of number of knots on the performance of penalized splines. Two algorithms for the automatic selection of the number of knots, myopic algorithm and full search algorithm, were proposed. Ruppert (2002) also described a Demmler-Reinsch type diagonalization for computing univariate and additive penalized spline, which is very useful for super-fast generalized cross-validation, while being not effective for smoothing splines since large number of knots.

The choices for the number and positioning of the knots are much less crucial than the smoothing parameter. Ruppert et al. (2003) introduced some model selection approaches, e.g., cross-validation (CV), generalized cross-validation (GCV), Mallows's C_p criterion. The optimal amount of smoothing in penalized spline regression was investigated in Wand (1999). In this article, a simple closed form approximation to the optimal smoothing parameter was derived. This approach was based on the mean average squared error (MASE), which is a mathematically measure of the global discrepancy between $\widehat{m}$ and m. It was shown to be a useful starting point for measuring the optimal amount of smoothing in penalized spline regression.

In nonparametric regression one can select the smoothing parameter by minimizing a Mean Squared Error (MSE) based criterion. For spline smoothing, the smooth estimation can be rewritten as a Linear Mixed Model. Then Maximum Likelihood (ML) theory can be applied to estimate the smoothing parameter as variance component. The relationship between spline smoothing and Mixed Models was discussed in Green and Silverman (1994), Brumback and Rice (1998) and Verbyla et al. (1999).

In Kauermann (2005), smoothing parameter selections for P-spline smoothing based on MSE minimization and REML estimation were compared. The results for MSE minimization method are similar to the results provided in Wand (1999). It was shown that REML-based smoothing parameter selection is asymptotically biased towards undersmoothing, i.e., this approach chooses a more complex model compared to the MSE method. The result accords with classical spline smoothing, however the asymptotic arguments are different.

Different smoother is rapidly becoming more popular, which is much easier to prove theoretical results. In this chapter, we will propose a new CV-based criterion for smoothing parameter selection, which has almost exact expression as the criterion (9) in Opsomer and Miller (2005) except that a penalized spline estimator is used to estimate the smooth function instead of a local polynomial estimator in Opsomer and Miller (2005).

Section 1.2.1 will give the definition of the spline estimator. Section 1.2.2 introduces the smoothing parameter selection method. In section 1.3 we state assumptions used in the theoretical derivations and our main theoretical results are described. In section 1.4, we report simulation results, which show how well the CV-based criterion works in practice.

1.2 Definition of the Estimator and Smoothing Parameter Selection

1.2.1 Definition of the Estimator

In survey sampling, the estimation of a finite population total,

$$t_y = \sum_{i \in U} y_i$$

is a problem in common. Where $U = \{1, 2, \ldots, N\}$ is a finite population with N identifiable elements and y_i is a response variable for the ith element. A sample of population elements $s \subset U$ is selected with probability $p(s)$. Let $\pi_i = \Pr(i \in s) = \sum_{s:i \in s} p(s) > 0$ denote the inclusion probability for element i, then the Horvitz-Thompson estimator (Horvitz and Thompson 1952) for t_y is

$$\widehat{t}_{y,\mathrm{HT}} = \sum_{s} \frac{y_i}{\pi_i}. \tag{1.1}$$

The variance of the Horvitz-Thompson estimator under the sampling design is

$$\mathrm{Var}\left(\widehat{t}_{y,\mathrm{HT}}\right) = \sum_{i \in U} \sum_{j \in U} \left(\pi_{ij} - \pi_i \pi_j\right) \frac{y_i}{\pi_i} \frac{y_j}{\pi_j}, \tag{1.2}$$

where $\pi_{ij} = \Pr(i \in s, j \in s)$, the joint inclusion probability for elements $i, j \in U$.

Suppose there is auxiliary information x_i available for all of U. Then we hope to improve estimation of t_y by using the auxiliary information. One approach to incorporate the auxiliary information is postulating a superpopulation model, say ξ, which describes the relationship between the response variable y and the auxiliary variable x.

Consider the superpopulation regression model

$$y_i = m(x_i) + \varepsilon_i, \tag{1.3}$$

where ε_i are independent random variables with mean zero and variance $v(x_i)$, $m(x_i)$ is a smooth function of x_i, and $v(x_i)$ is smooth and strictly positive. In order to introduce the estimator, we treat $\{(x_i, y_i) : i \in U\}$ as a realization from the superpopulation model (1.3). If the entire realization were observed, we could define a P-spline estimator for $m(\cdot)$ as follows:

$$m(x; \boldsymbol{\beta}) = \beta_0 + \beta_1 x + \ldots + \beta_q x^q + \sum_{k=1}^{K} \beta_{q+k} (x - \kappa_k)_+^q, \tag{1.4}$$

where $(t)_+^q = t^q$ if $t > 0$ and 0 otherwise, q is the degree of the spline, $\kappa_1 < \cdots \ldots < \kappa_K$ is a set of fixed knots and $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_{q+K})^T$ is the coefficient vector. Typically, q is kept fixed and low. If the number of knots K is sufficiently large, the class of functions $m(x; \boldsymbol{\beta})$ is very large and can approximate most smooth functions with a high degree of accuracy.

The population estimator for $\boldsymbol{\beta}$ is defined as the minimizer of

$$\sum_{i \in U} (y_i - m(x_i; \boldsymbol{\beta}))^2 + \alpha \sum_{k=1}^{K} \beta_{q+k}^2 \tag{1.5}$$

for some fixed constant $\alpha \geq 0$. The smoothness of the resulting fit depends on the value of α, with larger values corresponding to smoother fits.

Let $\boldsymbol{X}$ represent the matrix with rows $\boldsymbol{x}_i^T = \left\{1, x_i, \ldots, x_i^q, (x_i - \kappa_1)_+^q, \ldots, (x_i - \kappa_K)_+^q\right\}$ for $i \in U$, and let $\boldsymbol{Y}$ denote the column vector of response values y_i for $i \in U$. Define a diagonal matrix $\boldsymbol{A}_\alpha = \operatorname{diag}\{0, \ldots, 0, \alpha, \ldots, \alpha\}$, which has $q+1$ zeros followed by K penalty constant α. If the population U is fully observed, the penalized least squares estimator for the coefficient vector of (1.4) has the ridge-regression representation:

$$\boldsymbol{\beta}_U = \left(\boldsymbol{X}^T \boldsymbol{X} + \boldsymbol{A}_\alpha\right)^{-1} \boldsymbol{X}^T \boldsymbol{Y}. \tag{1.6}$$

Let $m_i = m(x_i; \boldsymbol{\beta}_U) \equiv \boldsymbol{x}_i^T \boldsymbol{\beta}_U$, $i \in U$ denote the P-spline fit obtained from this hypothetical population fit at x_i. If these fitted values are known, they can be incorporated into the survey estimation by constructing the difference estimator (Särndal et al. 1992, p. 221)

$$\widehat{t}_{y,\text{diff}} = \sum_U m_i + \sum_s \frac{y_i - m_i}{\pi_i}. \tag{1.7}$$

The difference estimator is design unbiased and its design variance is

$$\mathrm{Var}\left(\widehat{t}_{y,\mathrm{diff}}\right) = \sum_U \sum_U \left(\pi_{ij} - \pi_i \pi_j\right) \frac{y_i - m_i}{\pi_i} \frac{y_j - m_j}{\pi_j}. \tag{1.8}$$

Obviously, the efficiency of $\widehat{t}_{y,\mathrm{diff}}$ depends on how well the m_i approximates the variable y_i.

The estimator (1.7) is infeasible because the m_i cannot be calculated. However, given a sample s, the m_i in (1.7) can be replaced by sample-based estimators, denoted by $\widehat{m}_i$ and constructed as follows. Define the diagonal matrix of inverse inclusion probabilities $\boldsymbol{W} = \mathrm{diag}_{j\in U}\left\{1/\pi_j\right\}$ and its sample submatrix $\boldsymbol{W}_s = \mathrm{diag}_{j\in s}\left\{1/\pi_j\right\}$. Similarly, let $\boldsymbol{X}_s$ be the submatrix of $\boldsymbol{X}$ consisting of those rows for which $j \in s$ and $\boldsymbol{Y}_s$ denote the column vector of response values y_j for $j \in s$. For fixed α and under suitable regularity conditions, the π-weighted estimator

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{X}_s + \boldsymbol{A}_\alpha\right)^{-1} \boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{Y}_s = \boldsymbol{G}_\alpha \boldsymbol{Y}_s \tag{1.9}$$

is a design-consistent estimator of $\boldsymbol{\beta}_U$ in (1.6).

Define $\widehat{m}_i = m\left(x_i, \widehat{\boldsymbol{\beta}}\right) \equiv \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}$. Then the model-assisted P-spline estimator is defined as

$$\widehat{t}_{y,\mathrm{spl}} = \sum_U \widehat{m}_i + \sum_s \frac{y_i - \widehat{m}_i}{\pi_i}. \tag{1.10}$$

1.2.2 Smoothing Parameter Selection

Introducing the indicator function $I_i = 1$ if $i \in s$ and 0 otherwise, and the indicator vector $\boldsymbol{e}_i$ which is a zero vector except for an entry of one at position i, we can rewrite (1.10) as

$$\widehat{t}_{y,\mathrm{spl}} = \sum_{i\in s} \left\{ \pi_i^{-1} + \sum_{j\in U} \left(1 - I_j/\pi_j\right) \boldsymbol{x}_j^T \boldsymbol{G}_\alpha \boldsymbol{e}_i \right\} y_i \equiv \sum_s w_{i(s)} y_i,$$

which shows that $\widehat{t}_{y,\mathrm{spl}}$ is a linear estimator.

In Breidt et al. (2005), it was shown that this estimator is design consistent and asymptotically design unbiased. Asymptotically, the design mean squared error of $\widehat{t}_{y,\mathrm{spl}}$ is equivalent to the variance of the generalized difference estimator, given in (1.8),

$$\mathrm{MSE}_p\left(\widehat{t}_{y,\mathrm{spl}}\right) = \mathrm{E}_p\left(\widehat{t}_{y,\mathrm{spl}} - t_y\right)^2 \approx \sum_{i,j\in U} \left(\pi_{ij} - \pi_i \pi_j\right) \frac{y_i - m_i}{\pi_i} \frac{y_j - m_j}{\pi_j}. \tag{1.11}$$

Finally, Breidt et al. (2005) provides a design consistent and asymptotically design unbiased estimator of $\mathrm{MSE}_p\left(\widehat{t}_{y,\mathrm{spl}}\right)$, as

$$\widehat{V}\left(\widehat{t}_{y,\mathrm{spl}}\right) = \sum_{i,j\in s} \frac{\pi_{ij}-\pi_i\pi_j}{\pi_{ij}} \frac{y_i-\widehat{m}_i}{\pi_i} \frac{y_j-\widehat{m}_j}{\pi_j}. \tag{1.12}$$

The problem is that minimizing $\widehat{V}\left(\widehat{t}_{y,\mathrm{spl}}\right)$ does not lead to the minimizer of the MSE, since $\widehat{m}_i$ can be made to be close to y_i by letting $\alpha \to 0$.

In Opsomer and Miller (2005), $\widehat{V}\left(\widehat{t}_{y,\mathrm{spl}}\right)$ is modified so that it provides a more suitable criterion. Specifically, each estimator $\widehat{m}_i$ is replaced by the "leave-one-out" estimator $\widehat{m}_i^{(-)}$. This estimator is readily derived by defining a modified smoothing vector $\boldsymbol{w}'_{si}$ with elements

$$w'_{sij} = \begin{cases} \frac{w_{sij}}{1-w_{sii}} & \text{if } j \neq i \\ 0 & \text{if } i = j \end{cases},$$

where w_{sij} denotes the jth element of the vector $\boldsymbol{w}_{si} = \left(\boldsymbol{x}_i^T \boldsymbol{G}_\alpha\right)^T$, and set $\widehat{m}_i^{(-)} := \sum_{j\in s} w'_{sij} y_j$. The modification of $\widehat{V}\left(\widehat{t}_{y,\mathrm{spl}}\right)$ proposed to use is defined as

$$\widehat{V}_{\mathrm{CV}}\left(\widehat{t}_{y,\mathrm{spl}}\right) := \sum_{i,j\in s} \frac{\pi_{ij}-\pi_i\pi_j}{\pi_{ij}} \frac{y_i-\widehat{m}_i^{(-)}}{\pi_i} \frac{y_j-\widehat{m}_j^{(-)}}{\pi_j}. \tag{1.13}$$

We will refer to $\widehat{V}_{\mathrm{CV}}\left(\widehat{t}_{y,\mathrm{spl}}\right)$, denoted by $\widehat{V}_{\mathrm{CV}}(\alpha)$, as the CV criterion for smoothing parameter selection in function estimation. We will write $\widehat{\alpha}_{\mathrm{CV}}$ for the minimizer of $\widehat{V}_{\mathrm{CV}}(\alpha)$, and use it as an estimator of α_{opt}, the minimizer of $\mathrm{MSE}_p\left(\widehat{t}_{y,\mathrm{spl}}\right)$.

1.3 Theoretical Properties

In order to prove our theoretical results, we make the following technical assumptions (see Breidt and Opsomer (2000) and Breidt et al. (2005)). For simplicity, we will only consider the case with the sample size, denoted by n_N, fixed for each N, and also assume that $n_N \to \infty$. As above, q, K and $\{\kappa_k\}$ are fixed.

A1. (Sampling rate $n_N N^{-1}$). As $N \to \infty$, $n_N N^{-1} \to \pi \in (0,1)$.

A2. (Inclusion probabilities π_i and π_j). For all N, $\min_{i\in U} \pi_i \geq \lambda > 0$, $\min_{i,j\in U} \pi_{ij} \geq \lambda^* > 0$, $\lim_{N\to\infty} n_N \max_{i,j\in U, i\neq j} |\Delta_{ij}| < \infty$, where $\Delta_{ij} = \pi_{ij} - \pi_i\pi_j$.

A3. Let $\widehat{\boldsymbol{D}}_s = \frac{n_N}{N^2}\left(\boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{X}_s + \boldsymbol{A}_\alpha\right)$, and $\boldsymbol{D}_U = \frac{n_N}{N^2}\left(\boldsymbol{X}_U^T \boldsymbol{X}_U + \boldsymbol{A}_\alpha\right) = \mathrm{E}_p\left(\widehat{\boldsymbol{D}}_s\right)$. Assume $\widehat{\boldsymbol{D}}_s^{-1}$, $\boldsymbol{D}_U^{-1}$ exist for all $\alpha \in H_\alpha$, where H_α is an interval fixed between 0 and some constant C_α with $0 < C_\alpha < \infty$.

A4. $\max_{i\in U} |y_i| < C_y$, and $\max_{i\in U, j\in\{1,\ldots,p\}} |x_{ij}| < C_x$, where $p = 1+q+K$ is the dimension of $\boldsymbol{x}_i$, C_y and C_x are some positive constants.

A5. Additional assumptions involving higher-order inclusion probabilities:

$$\lim_{N\to\infty} \max_{(i,j,k)\in D_{3,N}} \left|\mathrm{E}_p\left[\left(I_i - \pi_i\right)\left(I_j I_k - \pi_{jk}\right)\right]\right| = 0$$
$$\lim_{N\to\infty} \max_{(i,j,k,l)\in D_{4,N}} \left|\mathrm{E}_p\left[\left(I_i I_j - \pi_{ij}\right)\left(I_k I_l - \pi_{kl}\right)\right]\right| = 0$$
$$\lim_{N\to\infty} n_N \max_{(i,j,k)\in D_{3,N}} \left|\mathrm{E}_p\left[\left(I_i - \pi_i\right)^2\left(I_j - \pi_j\right)\left(I_k - \pi_k\right)\right]\right| < \infty$$
$$\lim_{N\to\infty} n_N^2 \max_{(i,j,k,l)\in D_{4,N}} \left|\mathrm{E}_p\left[\left(I_i - \pi_i\right)\left(I_j - \pi_j\right)\left(I_k - \pi_k\right)\left(I_l - \pi_l\right)\right]\right| < \infty$$

where $D_{t,N}$ denotes the set of all distinct t-tuples from U.

The assumption 3 ensures that $\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{\beta}_U$ exist for all $\alpha \in H_\alpha$. The assumption depends on the knots, the penalty constant α, and the distribution of the $\boldsymbol{x}_i$.

The following results establish design consistency of variance estimator of $\widehat{t}_{y,\mathrm{spl}}$.

Theorem 1.3.1. *Let* $w_{sii} = \frac{n_N}{N^2}\boldsymbol{x}_i^T \widehat{\boldsymbol{D}}_s^{-1}\boldsymbol{x}_i/\pi_i$, *and assumptions A1-A5 hold. Then, the auxiliary population* $\{x_j\}_{j\in U}$, *error population* $\{\varepsilon_j\}_{j\in U}$, *and sample* s *are such that*

$$\sup_{\alpha\in H_\alpha}\left|\widehat{V}_{\mathrm{CV}}\left(\widehat{t}_{y,\mathrm{spl}}\right) - \mathrm{Var}\left(\widehat{t}_{y,\mathrm{diff}}\right)\right| = o_p\left(\frac{N^2}{n_N}\right),$$

where $\mathrm{Var}\left(\widehat{t}_{y,\mathrm{diff}}\right) = \sum\sum_U \Delta_{ij}\frac{y_i - m_i}{\pi_i}\frac{y_j - m_j}{\pi_j}$.

Proof of Theorem 1.3.1: We write the expression as follows

$$\begin{aligned}
\widehat{V}_{\mathrm{CV}}\left(\widehat{t}_{y,\mathrm{spl}}\right) &= \sum_{i\in s}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\frac{y_i - \widehat{m}_i}{\pi_i}\frac{y_j - \widehat{m}_j}{\pi_j}\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}} \\
&= \sum_{i\in s}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\frac{y_i - m_i}{\pi_i}\frac{y_j - m_j}{\pi_j}\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}} \\
&\quad + \sum_{i\in s}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\frac{m_i - \widehat{m}_i}{\pi_i}\frac{m_j - \widehat{m}_j}{\pi_j}\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}} \\
&\quad + 2\sum_{i\in s}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\frac{y_i - m_i}{\pi_i}\frac{m_j - \widehat{m}_j}{\pi_j}\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}} \\
&= V_1 + V_2 + 2V_3.
\end{aligned}$$

In this expression,

$$
\begin{aligned}
V_1 &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}}\\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}}\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\\
&= V_{11}+V_{12}.
\end{aligned}
$$

And,

$$
\begin{aligned}
V_{11} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}\\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}\left(\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}}-1\right)\\
&= \operatorname{Var}\left(\widehat{t}_{y,\text{diff}}\right)+V_{11}^*.
\end{aligned}
$$

Suppose assumptions A1, A2 hold, from Lemma A.1.3 and A.1.5, we can show that

$$
\begin{aligned}
\sup_{\alpha\in H_\alpha}\left|V_{11}^*\right| &\le \sup_{\alpha\in H_\alpha}\sum_{i\in U}\left|\pi_i\left(1-\pi_i\right)\right|\frac{\left(y_i-m_i\right)^2}{\pi_i^2}\left|g_{ii}\left(\widehat{\boldsymbol{D}}_s\right)\right|+\\
&\quad \sup_{\alpha\in H_\alpha}\sum_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\left|\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}\right|\left|g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\right|\\
&\le \frac{1}{\lambda^2}\sum_{i\in U}\left\{\sup_{\alpha\in H_\alpha}\max_{i\in U}\left|y_i-m_i\right|\right\}^2\sup_{\alpha\in H_\alpha}\max_{i,j\in U}\left|g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\right|\\
&\quad+\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\left\{\sup_{\alpha\in H_\alpha}\max_{i\in U}\left|y_i-m_i\right|\right\}^2\sup_{\alpha\in H_\alpha}\max_{i,j\in U}\left|g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\right|\\
&= O\left(1\right)+O\left(\frac{N}{n_N}\right)\\
&= O\left(\frac{N}{n_N}\right),
\end{aligned}
$$

as the fact that $\max_{i\in U}\left|\Delta_{ii}\right| = \max_{i\in U}\left|\pi_i\left(1-\pi_i\right)\right| \le 1$. Rewrite V_{12} as follows,

$$
\begin{aligned}
V_{12} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}\left(\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}}-1\right)\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\\
&= V_{121}+V_{122}.
\end{aligned}
$$

Where, $\mathrm{E}_p\left[V_{121}\right]=0$, and

$$
\begin{aligned}
\mathrm{Var}_p\left[V_{121}\right] &= E_p\left[V_{121}^2\right] \\
&= \sum_{i,j,k,l\in U}\Delta_{ij}\Delta_{kl}\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}\frac{y_k-m_k}{\pi_k}\frac{y_l-m_l}{\pi_l}\mathrm{E}_p\left[\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\frac{I_kI_l}{\pi_{kl}}-1\right)\right]
\end{aligned}
$$

Then, by assumptions A1, A2, A5 and Lemma A.1.3,

$$
\begin{aligned}
\sup_{\alpha\in H_\alpha}\left|\mathrm{Var}_p\left[V_{121}\right]\right| &\leq \frac{1}{\lambda^4}\sum_{i,j,k,l\in U}\max_{i,j\in U}\left|\Delta_{ij}\right|\max_{k,l\in U}\left|\Delta_{kl}\right|\left\{\sup_{\alpha\in H_\alpha}\max_{i\in U}\left|y_i-m_i\right|\right\}^4 \\
&\quad\times\max_{i,j,k,l\in U}\left|E_p\left[\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\frac{I_kI_l}{\pi_{kl}}-1\right)\right]\right| \\
&= \frac{2}{\lambda^5\lambda^*}\sum_{(i,j,k)\in D_{3,N}}O\left(\frac{1}{n_N}\right)\max_{(i,j,k)\in D_{3,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_k-\pi_k\right)\right]\right| \\
&\quad+\frac{4}{\lambda^4{\lambda^*}^2}\sum_{(i,j,k)\in D_{3,N}}O\left(\frac{1}{n_N^2}\right)\max_{(i,j,k)\in D_{3,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_iI_k-\pi_{ik}\right)\right]\right| \\
&\quad+\frac{4}{\lambda^5\lambda^*}\sum_{(i,j)\in D_{2,N}}O\left(\frac{1}{n_N}\right)\max_{(i,j)\in D_{2,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_i-\pi_i\right)\right]\right| \\
&\quad+\frac{1}{\lambda^6}\sum_{i\in U}O(1)\max_{i\in U}\left|E_p\left[\left(I_i-\pi_i\right)^2\right]\right| \\
&\quad+\frac{1}{\lambda^4{\lambda^*}^2}\sum_{(i,j,k,l)\in D_{4,N}}O\left(\frac{1}{n_N^2}\right)\max_{(i,j,k,l)\in D_{4,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_kI_l-\pi_{kl}\right)\right]\right| \\
&= o\left(\frac{N^3}{n_N}\right)+O\left(\frac{N^3}{n_N^2}\right)+O\left(\frac{N^2}{n_N}\right)+O(N)+o\left(\frac{N^4}{n_N^2}\right) \\
&= o\left(\frac{N^4}{n_N^2}\right),
\end{aligned}
$$

which implies that $\sup_{\alpha\in H_\alpha}\left|V_{121}\right|=o_p\left(\frac{N^2}{n_N}\right)$.

By assumption A2,

$$
\begin{aligned}
\max_{i\in U}\left|\frac{I_i}{\pi_i}-1\right| &= \max_{i\in U}\left(1,\frac{1}{\pi_i}-1\right) \\
&\leq \frac{1}{\lambda}, \qquad (1.14)
\end{aligned}
$$

and,

$$
\begin{aligned}
\max_{(i,j)\in D_{2,N}}\left|\frac{I_iI_j}{\pi_{ij}}-1\right| &= \max_{i\in U}\left(1,\frac{1}{\pi_{ij}}-1\right) \\
&\leq \frac{1}{\lambda^*}, \qquad (1.15)
\end{aligned}
$$

then V_{122} should have the same order as V_{11}^*. Therefore, by assumptions A1, A2, Lemma A.1.3 and A.1.5, it can be shown that

$$\begin{aligned}
\sup_{\alpha\in H_\alpha} |V_{122}| &\leq \sum_{i\in U} \frac{1}{\lambda^3} \left\{ \sup_{\alpha\in H_\alpha} \max_{i\in U} |y_i - m_i| \right\}^2 \sup_{\alpha\in H_\alpha} \max_{i,j\in U} \left| g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\right| \\
&\quad + \sum_{(i,j)\in D_{2,N}} \frac{1}{\lambda^2\lambda^*} \max_{(i,j)\in D_{2,N}} |\Delta_{ij}| \left\{ \sup_{\alpha\in H_\alpha} \max_{i\in U} |y_i - m_i| \right\}^2 \sup_{\alpha\in H_\alpha} \max_{i,j\in U} \left| g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\right| \\
&= O(1) + O\left(\frac{N}{n_N}\right) \\
&= O\left(\frac{N}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U} |\Delta_{ii}| = \max_{i\in U} |\pi_i (1-\pi_i)| \leq 1$. Thus, it follows that

$$\begin{aligned}
\sup_{\alpha\in H_\alpha} |V_{12}| &\leq \sup_{\alpha\in H_\alpha} |V_{121}| + \sup_{\alpha\in H_\alpha} |V_{122}| \\
&= o_p\left(\frac{N^2}{n_N}\right) + O\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}$$

Next, we will show that

$$\sup_{\alpha\in H_\alpha} |V_2| = o_p\left(\frac{N^2}{n_N}\right).$$

Note

$$\begin{aligned}
V_2 &= \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{m_i - \widehat{m}_i}{\pi_i} \frac{m_j - \widehat{m}_j}{\pi_j} \frac{1}{1-w_{sii}} \frac{1}{1-w_{sjj}} \\
&\quad + \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{m_i - \widehat{m}_i}{\pi_i} \frac{m_j - \widehat{m}_j}{\pi_j} \frac{1}{1-w_{sii}} \frac{1}{1-w_{sjj}} \left(\frac{I_i I_j}{\pi_{ij}} - 1\right) \\
&= V_{21} + V_{22}.
\end{aligned}$$

And,

$$\begin{aligned}
V_{21} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{m_i-\widehat{m}_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j} \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{m_i-\widehat{m}_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j}\left(\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}}-1\right) \\
&= \left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right) \\
&\quad+\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right) \\
&= V_{211}+V_{212}.
\end{aligned}$$

Suppose assumptions A1, A2 and A4 hold for all $\alpha \in H_\alpha$, by Lemma A.1.1, it can be shown that

$$\begin{aligned}
\sup_{\alpha\in H_\alpha}|V_{211}| &\leq \frac{1}{\lambda^2}\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right|^T\sum_{i\in U}\left|\pi_i(1-\pi_i)\boldsymbol{x}_i\boldsymbol{x}_i^T\right|\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right| \\
&\quad+\frac{1}{\lambda^2}\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right|^T\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}|\Delta_{ij}|\left|\boldsymbol{x}_i\boldsymbol{x}_j^T\right|\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right| \\
&= o_p(1)\,O(N)\,o_p(1)+o_p(1)\,O\left(\frac{N^2}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N^2}{n_N}\right),
\end{aligned}$$

and by Lemma A.1.5

$$\begin{aligned}
\sup_{\alpha\in H_\alpha}|V_{212}| &\leq \frac{1}{\lambda^2}\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right|^T\sum_{i\in U}\left|\pi_i(1-\pi_i)\boldsymbol{x}_i\boldsymbol{x}_i^T\right|\sup_{\alpha\in H_\alpha}\max_{i,j\in U}\left|g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\right|\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right| \\
&\quad+\frac{1}{\lambda^2}\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right|^T\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}|\Delta_{ij}|\left|\boldsymbol{x}_i\boldsymbol{x}_j^T\right| \\
&\quad\times\sup_{\alpha\in H_\alpha}\max_{i,j\in U}\left|g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\right|\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right| \\
&= o_p(1)\,O(1)\,o_p(1)+o_p(1)\,O\left(\frac{N}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U}|\Delta_{ii}| = \max_{i\in U}|\pi_i(1-\pi_i)| \le 1$. Then, it follows that

$$\begin{aligned}\sup_{\alpha\in H_\alpha}|V_{21}| &\le \sup_{\alpha\in H_\alpha}|V_{211}| + \sup_{\alpha\in H_\alpha}|V_{212}| \\ &= o_p\left(\frac{N^2}{n_N}\right) + o_p\left(\frac{N}{n_N}\right) \\ &= o_p\left(\frac{N^2}{n_N}\right).\end{aligned}$$

Rewrite V_{22} as

$$\begin{aligned}V_{22} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{m_i-\widehat{m}_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j}\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\ &\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{m_i-\widehat{m}_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j}\left(\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}}-1\right)\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\ &= \left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right) \\ &\quad+\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right) \\ &= V_{221}+V_{222}.\end{aligned}$$

By (1.14) and (1.15), V_{22} should have the same order as V_{21}. It follows that

$$\sup_{\alpha\in H_\alpha}|V_{22}| = o_p\left(\frac{N^2}{n_N}\right).$$

Therefore, it can be shown that

$$\begin{aligned}\sup_{\alpha\in H_\alpha}|V_2| &\le \sup_{\alpha\in H_\alpha}|V_{21}| + \sup_{\alpha\in H_\alpha}|V_{22}| \\ &= o_p\left(\frac{N^2}{n_N}\right) + o_p\left(\frac{N^2}{n_N}\right) \\ &= o_p\left(\frac{N^2}{n_N}\right).\end{aligned}$$

Finally, we will show that

$$\sup_{\alpha\in H_\alpha}|V_3| = o_p\left(\frac{N^2}{n_N}\right).$$

Note

$$
\begin{aligned}
V_3 &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j}\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}} \\
&\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j}\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}}\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\
&= V_{31}+V_{32},
\end{aligned}
$$

where

$$
\begin{aligned}
V_{31} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j} \\
&\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j}\left(\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}}-1\right) \\
&= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\boldsymbol{\beta}_U-\widehat{\boldsymbol{\beta}}\right) \\
&\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\left(\boldsymbol{\beta}_U-\widehat{\boldsymbol{\beta}}\right) \\
&= V_{311}+V_{312}.
\end{aligned}
$$

Suppose assumptions A1, A2 and A4 hold, from Lemma A.1.1, A.1.3 and A.1.5 we can show that

$$
\begin{aligned}
\sup_{\alpha\in H_\alpha}|V_{311}| &= \sup_{\alpha\in H_\alpha}\left|\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\boldsymbol{\beta}_U-\widehat{\boldsymbol{\beta}}\right)\right| \\
&\leq \frac{1}{\lambda^2}\sum_{i\in U}\sup_{\alpha\in H_\alpha}\max_{i\in U}|y_i-m_i|\max_{i\in U}|\boldsymbol{x}_i|^T\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right| \\
&\quad +\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}|\Delta_{ij}|\sup_{\alpha\in H_\alpha}\max_{i\in U}|y_i-m_i|\max_{j\in U}|\boldsymbol{x}_j|^T\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right| \\
&= O(N)\,o_p(1)+O\left(\frac{N^2}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N^2}{n_N}\right),
\end{aligned}
$$

and,

$$
\begin{aligned}
\sup_{\alpha\in H_\alpha}|V_{312}| &= \sup_{\alpha\in H_\alpha}\left|\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\left(\boldsymbol{\beta}_U-\widehat{\boldsymbol{\beta}}\right)\right| \\
&\leq \frac{1}{\lambda^2}\sum_{i\in U}\sup_{\alpha\in H_\alpha}\max_{i\in U}|y_i-m_i|\max_{i\in U}|\boldsymbol{x}_i|^T\sup_{\alpha\in H_\alpha}\max_{i,j\in U}\left|g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\right|\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right| \\
&\quad +\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}|\Delta_{ij}|\sup_{\alpha\in H_\alpha}\max_{i\in U}|y_i-m_i|\max_{j\in U}|\boldsymbol{x}_j|^T \\
&\quad \times\sup_{\alpha\in H_\alpha}\max_{i,j\in U}\left|g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\right|\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right| \\
&= O(1)\,o_p(1)+O\left(\frac{N}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N}{n_N}\right),
\end{aligned}
$$

as the fact that $\max_{i\in U}|\Delta_{ii}| = \max_{i\in U}|\pi_i(1-\pi_i)| \leq 1$. Thereby,

$$
\begin{aligned}
\sup_{\alpha\in H_\alpha}|V_{31}| &\leq \sup_{\alpha\in H_\alpha}|V_{311}|+\sup_{\alpha\in H_\alpha}|V_{312}| \\
&= o_p\left(\frac{N^2}{n_N}\right)+o_p\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}
$$

Write V_{32} as follows

$$
\begin{aligned}
V_{32} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j}\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\
&\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{m_j-\widehat{m}_j}{\pi_j}\left(\frac{1}{1-w_{sii}}\frac{1}{1-w_{sjj}}-1\right)\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\
&= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\boldsymbol{\beta}_U-\widehat{\boldsymbol{\beta}}\right) \\
&\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}\left(\widehat{\boldsymbol{D}}_s\right)\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\boldsymbol{\beta}_U-\widehat{\boldsymbol{\beta}}\right) \\
&= V_{321}+V_{322}.
\end{aligned}
$$

Similarly, from (1.14) and (1.15), V_{32} should have the same order as V_{31}. It follows that

$$\begin{aligned}\sup_{\alpha\in H_\alpha}|V_3| &\le \sup_{\alpha\in H_\alpha}|V_{31}|+\sup_{\alpha\in H_\alpha}|V_{32}|\\ &= o_p\left(\frac{N^2}{n_N}\right)+o_p\left(\frac{N^2}{n_N}\right)\\ &= o_p\left(\frac{N^2}{n_N}\right).\end{aligned}$$

Therefore,

$$\begin{aligned}\sup_{\alpha\in H_\alpha}\left|\widehat{V}_{\text{CV}}\left(\widehat{t}_{y,\text{spl}}\right)-\text{Var}\left(\widehat{t}_{y,\text{diff}}\right)\right| &\le \sup_{\alpha\in H_\alpha}|V_{11}^*|+\sup_{\alpha\in H_\alpha}|V_{12}|+\sup_{\alpha\in H_\alpha}|V_2|+2\sup_{\alpha\in H_\alpha}|V_3|\\ &= O\left(\frac{N}{n_N}\right)+o_p\left(\frac{N^2}{n_N}\right)+o_p\left(\frac{N^2}{n_N}\right)+2o_p\left(\frac{N^2}{n_N}\right)\\ &= o_p\left(\frac{N^2}{n_N}\right).\end{aligned}$$

Thus, the result follows.

□

Theorem 1.3.2. *Let assumptions A1-A5 hold. Then, the auxiliary population* $\{x_j\}_{j\in U}$*, error population* $\{\varepsilon_j\}_{j\in U}$*, and sample* s *are such that*

$$\lim_{N\to\infty}\sup_{\alpha\in H_\alpha}\frac{1}{N}\left|\text{MSE}_p\left(\widehat{t}_{y,\text{spl}}\right)-\text{Var}\left(\widehat{t}_{y,\text{diff}}\right)\right|=0,$$

where $\text{Var}\left(\widehat{t}_{y,\text{diff}}\right)=\sum\sum_U\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}$.

Proof of Theorem 1.3.2: Let $a_N=\sum_U(y_i-m_i)\left(\frac{I_i}{\pi_i}-1\right)$, $b_N=\sum_U(m_i-\widehat{m}_i)\left(\frac{I_i}{\pi_i}-1\right)$. Then,

$$\begin{aligned}\text{E}_p\left[a_N^2\right] &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-m_i}{\pi_i}\frac{y_j-m_j}{\pi_j}\\ &= \text{Var}\left(\widehat{t}_{y,\text{diff}}\right).\end{aligned}$$

And under assumptions A1, A2, by Lemma A.1.3,

$$\begin{aligned}
\mathrm{E}_p\left[a_N^2\right] &\leq \frac{1}{\lambda^2}\sum_{i\in U}\max_{i\in U}\left|\pi_i\left(1-\pi_i\right)\right|\left\{\sup_{\alpha\in H_\alpha}\max_{i\in U}\left|y_i-m_i\right|\right\}^2 \\
&\quad +\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\left\{\sup_{\alpha\in H_\alpha}\max_{i\in U}\left|y_i-m_i\right|\right\}^2 \\
&= O\left(N\right)+O\left(\frac{N^2}{n_N}\right) \\
&= O\left(\frac{N^2}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U}\left|\Delta_{ii}\right| = \max_{i\in U}\left|\pi_i\left(1-\pi_i\right)\right| \leq 1$. Then it follows that

$$\sup_{\alpha\in H_\alpha}\frac{1}{N}\mathrm{E}_p\left[a_N^2\right] = O\left(\frac{N}{n_N}\right).$$

Let b_k denote the kth element of $\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_U$, from assumptions A1, A2, A4 and A5, by Lemma A.1.2 we can show that

$$
\begin{aligned}
\mathrm{E}_p\left[b_N^2\right] &= \mathrm{E}_p\left[\sum_{(i,j)\in U}\sum_{k,l\in\{1,\ldots,p\}} x_{ik}x_{jl}b_kb_l\left(\frac{I_i}{\pi_i}-1\right)\left(\frac{I_j}{\pi_j}-1\right)\right] \\
&\leq \sum_{k,l\in\{1,\ldots,p\}}\sqrt{\mathrm{E}_p\left[b_k^2b_l^2\right]}\sqrt{\mathrm{E}_p\left[\left\{\sum_{(i,j)\in U}x_{ik}x_{jl}\left(\frac{I_i}{\pi_i}-1\right)\left(\frac{I_j}{\pi_j}-1\right)\right\}^2\right]} \\
&\leq \sum_{k,l\in\{1,\ldots,p\}}\frac{1}{\lambda^2}\sqrt{\sup_{\alpha\in H_\alpha}\max_{k\in\{1,\ldots,p\}}E_p\left[b_k^4\right]}\left\{\sqrt{\sum_{i\in U}\left\{\max_{i\in U,k\in\{1,\ldots,p\}}|x_{ik}|\right\}^4\max_{i\in U}\left|\mathrm{E}_p\left[(I_i-\pi_i)^4\right]\right|}\right. \\
&\quad+\sqrt{\sum_{(i,j)\in D_{2,N}}\left\{\max_{i\in U,k\in\{1,\ldots,p\}}|x_{ik}|\right\}^4\max_{(i,j)\in D_{2,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^2(I_j-\pi_j)^2\right]\right|} \\
&\quad+\sqrt{\sum_{(i,j)\in D_{2,N}}\left\{\max_{i\in U,k\in\{1,\ldots,p\}}|x_{ik}|\right\}^4\max_{(i,j)\in D_{2,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^3(I_j-\pi_j)\right]\right|} \\
&\quad+\sqrt{\sum_{(i,j,i')\in D_{3,N}}\left\{\max_{i\in U,k\in\{1,\ldots,p\}}|x_{ik}|\right\}^4\max_{(i,j,i')\in D_{3,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^2(I_j-\pi_j)(I_{i'}-\pi_{i'})\right]\right|} \\
&\quad+\sqrt{\sum_{(i,j,i',j')\in D_{4,N}}\left\{\max_{i\in U,k\in\{1,\ldots,p\}}|x_{ik}|\right\}^4} \\
&\quad\left.\times\sqrt{\max_{(i,j,i',j')\in D_{4,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)(I_j-\pi_j)(I_{i'}-\pi_{i'})(I_{j'}-\pi_{j'})\right]\right|}\right\} \\
&= O\left(\frac{1}{n_N}\right)O\left(\sqrt{N}\right)+O\left(\frac{1}{n_N}\right)O(N)+O\left(\frac{1}{n_N}\right)O(N)+O\left(\frac{1}{n_N}\right)O\left(\frac{N^{3/2}}{\sqrt{n_N}}\right) \\
&\quad+O\left(\frac{1}{n_N}\right)O\left(\frac{N^2}{n_N}\right) \\
&= o\left(\frac{N^2}{n_N}\right),
\end{aligned}
$$

which implies that

$$
\sup_{\alpha\in H_\alpha}\frac{1}{N}\mathrm{E}_p\left[b_N^2\right] = o\left(\frac{N}{n_N}\right),
$$

and,

$$
\begin{aligned}
\sup_{\alpha\in H_\alpha}\frac{1}{N}\mathrm{E}_p\left[|a_Nb_N|\right] &\leq \sqrt{\sup_{\alpha\in H_\alpha}\frac{1}{N}\mathrm{E}_p\left[a_N^2\right]\sup_{\alpha\in H_\alpha}\frac{1}{N}\mathrm{E}_p\left[b_N^2\right]} \\
&= \sqrt{O\left(\frac{N}{n_N}\right)o\left(\frac{N}{n_N}\right)} \\
&= o\left(\frac{N}{n_N}\right).
\end{aligned}
$$

Rewrite $\mathrm{MSE}_p\left(\widehat{t}_{y,\mathrm{spl}}\right)$ as follows,

$$\begin{aligned}
\mathrm{MSE}_p\left(\widehat{t}_{y,\mathrm{spl}}\right) &= \mathrm{E}_p\left[\left(\widehat{t}_{y,\mathrm{spl}} - t_y\right)^2\right] \\
&= \mathrm{E}_p\left[\left(\sum_{i\in U}\widehat{m}_i + \sum_{i\in s}\frac{y_i - \widehat{m}_i}{\pi_i} - \sum_{i\in U} y_i\right)^2\right] \\
&= \mathrm{E}_p\left[\left(\sum_{i\in U}\left(y_i - \widehat{m}_i\right)\left(\frac{I_i}{\pi_i} - 1\right)\right)^2\right] \\
&= \mathrm{E}_p\left[\left(\sum_{i\in U}\left(y_i - m_i + m_i - \widehat{m}_i\right)\left(\frac{I_i}{\pi_i} - 1\right)\right)^2\right] \\
&= \mathrm{E}_p\left[\left(a_N + b_N\right)^2\right] \\
&= \mathrm{E}_p\left[a_N^2\right] + \mathrm{E}_p\left[b_N^2\right] + 2\mathrm{E}_p\left[a_N b_N\right].
\end{aligned}$$

Then,

$$\begin{aligned}
\sup_{\alpha\in H_\alpha}\frac{1}{N}\left|\mathrm{MSE}_p\left(\widehat{t}_{y,\mathrm{spl}}\right) - \mathrm{Var}\left(\widehat{t}_{y,\mathrm{diff}}\right)\right| &= \sup_{\alpha\in H_\alpha}\frac{1}{N}\left|\mathrm{MSE}_p\left(\widehat{t}_{y,\mathrm{spl}}\right) - \mathrm{E}_p\left[a_N^2\right]\right| \\
&\leq \sup_{\alpha\in H_\alpha}\frac{1}{N}\mathrm{E}_p\left[b_N^2\right] + 2\sup_{\alpha\in H_\alpha}\frac{1}{N}\left|\mathrm{E}_p\left[a_N b_N\right]\right| \\
&\leq \sup_{\alpha\in H_\alpha}\frac{1}{N}\mathrm{E}_p\left[b_N^2\right] + 2\sup_{\alpha\in H_\alpha}\frac{1}{N}\mathrm{E}_p\left[\left|a_N b_N\right|\right] \\
&= o\left(\frac{N}{n_N}\right) + o\left(\frac{N}{n_N}\right) \\
&= o\left(\frac{N}{n_N}\right).
\end{aligned}$$

Therefore the result follows.

□

Corollary 1.3.1. *Let assumptions A1-A5 hold. Then, the auxiliary population $\{x_j\}_{j\in U}$, error population $\{\varepsilon_j\}_{j\in U}$, and sample s are such that*

$$\sup_{\alpha\in H_\alpha}\left|\widehat{V}_{\mathrm{CV}}\left(\widehat{t}_{y,\mathrm{spl}}\right) - \mathrm{MSE}_p\left(\widehat{t}_{y,\mathrm{spl}}\right)\right| = o_p\left(\frac{N^2}{n_N}\right).$$

Thereby the theory derived above for the P-spline estimator shows that it is possible to use $\widehat{V}_{\mathrm{CV}}\left(\widehat{t}_{y,\mathrm{spl}}\right)$, denoted by $\widehat{V}_{\mathrm{CV}}(\alpha)$, as an asymptotically equivalent criterion to $\mathrm{MSE}_p\left(\widehat{t}_{y,\mathrm{spl}}\right)$, denoted by $\mathrm{MSE}_p(\alpha)$, for selecting an optimal smoothing parameter α_{opt}. In section 1.4, we will evaluate how well this selection criterion works.

1.4 Simulation Results

In this section, we follow Opsomer and Miller (2005) in the design of a simulation study. A random population of $N = 1000$ values of x is generated from the uniform distribution on $[0, 1]$, and 1000 values for the errors ε are drawn from $\mathcal{N}(0, 1)$. This one error population is used for all simulations, up to multiplication by σ. Eight populations of y are generated as follows:

$$y_{il} = m_l\left(x_i\right) + \varepsilon_i \qquad 1 \le i \le 1000, \qquad 1 \le l \le 8,$$

where $\{m_l\}_{l=1}^{8}$ are predefined functions given in the Table 1.1. The finite population quantities of interest are $t_y = \sum_{i=1}^{1000} y_{il}$ for each l.

Population	Expression
1.Linear	$m_1\left(x\right) = 2x$
2.Quadratic	$m_2\left(x\right) = 1 + 2\left(x - 0.5\right)^2$
3.Bump	$m_3\left(x\right) = 2x + \exp\left(-200\left(x - 0.5\right)^2\right)$
4.Jump	$m_4\left(x\right) = 2xI_{\{x \le 0.65\}} + 0.65I_{\{x > 0.65\}}$
5.Normal CDF	$m_5\left(x\right) = \Phi\left(1.5 - 2x\right)$, where Φ is the standard normal cdf
6.Exponential	$m_6\left(x\right) = \exp\left(-8x\right)$
7.Slow sine	$m_7\left(x\right) = 2 + \sin\left(2\pi x\right)$
8.Fast sine	$m_8\left(x\right) = 2 + \sin\left(8\pi x\right)$

Table 1.1 Eight population mean functions.

The samples are drawn by one of two designs, simple random sampling without replacement (SI) or stratified simple random sampling without replacement (STSI). For each simulation run, $M = 1000$ samples are drawn from $\{(x_i, y_i)\}$. For each sample, we compute the estimator $\widehat{t}_{y,\text{spl}}$ in equation (1.10) for α_{opt} and $\widehat{\alpha}_{\text{CV}}$. Referring to Opsomer and Miller (2005), the optimal smoothing parameters α_{opt} for each population are not sample-based. We compute them by minimizing a simulation-based approximation to the function $\text{MSE}_p\left(\alpha\right)$, which is constructed by simulating repeated samples from these populations for a grid of smoothing parameters over the interval $[0.0001, 10]$, and finding the functions $\text{MSE}_p\left(\alpha\right)$ by averaging over these simulations.

For each sample, the smoothing parameter $\widehat{\alpha}_{\text{CV}}$ is sought through a search algorithm implemented in R, which uses expression:

$$\widehat{V}_{\text{CV}}(\alpha) := \sum_{i,j\in s} \frac{\pi_{ij} - \pi_i\pi_j}{\pi_{ij}} \frac{y_i - \widehat{m}_i^{(-)}}{\pi_i} \frac{y_j - \widehat{m}_j^{(-)}}{\pi_j},$$

the same expression as (1.13). A simulation run is determined by sample size n, error variance σ^2, and degree of the spline regression q. For the design of simple random sampling without replacement, simulations are done for $n \in \{100, 200, 500\}$, $\sigma^2 \in \{0.01, 0.16\}$, $q \in \{1\}$. The design of stratified simple random sampling without replacement uses 4 strata with each stratum containing 250 elements, and the stratification of the strata is based on a random variable z_i and ratio r. First, we generate v_i from a standard normal distribution $\mathcal{N}(0, \sigma_v^2)$ with σ_v^2 satisfying $r = \frac{\sigma^2}{\sigma^2+\sigma_v^2}$. Then z_i's are derived as follows:

$$z_i = \begin{cases} v_i + \varepsilon_i & \text{if } 0 < r < 1 \\ v_i & \text{if } r = 0, \text{ where } \sigma_v^2 = 1 \\ \varepsilon_i & \text{if } r = 1 \end{cases}.$$

After sorting by z_i $(i = 1, \ldots, N)$, the population is separated into 4 strata with boundaries given by equally-spaced quantiles of z. Then, simulations are conducted with the stratum sample sizes $\{(15, 20, 30, 35), (30, 40, 60, 70), (75, 100, 150, 175)\}$, $r \in \{0, 0.25, 0.5, 0.75, 1\}$, $\sigma^2 \in \{0.01, 0.16\}$, and $q = 1$. Thus, the strata have different sampling rates with the inclusion probability correlated with the model error.

As mentioned in Breidt et al. (2005), for m_1 and m_2, the models are polynomial; the remaining mean functions are representatives with various departures from the polynomial model. The mean function m_3 is mostly linear over its range, except that there is a 'bump' for a small portion of the range of x_k. Function m_4 is not a smooth function. The sigmoidal function m_5 is the cumulative distribution function, and m_6 is an exponential curve. The function m_7 is a sinusoid completing one full cycle on $[0, 1]$, while m_8 completes four full cycles.

Since the true α_{opt} in the case with model correctly specified increases to infinity, for simplicity, we restrict the range for searching the minimums of α_{opt} and $\widehat{\alpha}_{\text{CV}}$ within $(0, 10]$.

Population	σ	n	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
1.Linear	0.1	100	8.864	90.890	10.000	89.994
	0.1	200	9.121	33.828	10.000	33.800
	0.1	500	9.499	9.748	10.000	9.734
	0.4	100	8.864	1454.234	10.000	1439.899
	0.4	200	9.121	541.251	10.000	540.804
	0.4	500	9.499	155.972	10.000	155.738
2.Quadratic	0.1	100	1.668	92.636	1.385	91.601
	0.1	200	1.697	34.125	1.385	33.981
	0.1	500	1.102	9.749	0.774	9.742
	0.4	100	7.534	1464.405	7.925	1446.792
	0.4	200	7.956	543.012	4.977	541.553
	0.4	500	6.030	156.101	2.783	155.799
3.Bump	0.1	100	0.004	107.197	0.003	105.423
	0.1	200	0.004	37.898	0.003	37.739
	0.1	500	0.003	10.044	0.003	10.022
	0.4	100	0.173	1603.633	0.027	1547.149
	0.4	200	0.037	582.408	0.027	581.066
	0.4	500	0.016	159.168	0.015	158.539
4.Jump	0.1	100	0.005	124.059	0.003	120.596
	0.1	200	0.001	43.385	0.002	42.829
	0.1	500	0.000	10.987	0.001	10.921
	0.4	100	1.698	1555.201	0.085	1525.090
	0.4	200	0.368	578.733	0.171	567.463
	0.4	500	0.023	158.539	0.048	156.861
5.Normal CDF	0.1	100	8.290	90.999	10.000	90.057
	0.1	200	8.503	33.892	10.000	33.784
	0.1	500	7.753	9.723	10.000	9.691
	0.4	100	8.858	1455.499	10.000	1439.309
	0.4	200	9.161	541.070	10.000	540.494
	0.4	500	9.528	155.744	10.000	155.516
6.Exponential	0.1	100	0.158	95.709	0.048	94.708
	0.1	200	0.111	35.334	0.152	34.845
	0.1	500	0.043	9.914	0.027	9.889
	0.4	100	4.306	1486.481	1.963	1472.820
	0.4	200	2.908	549.327	1.556	544.119
	0.4	500	1.069	157.463	0.215	156.945
7.Slow sine	0.1	100	0.094	94.886	0.076	93.751
	0.1	200	0.103	35.249	0.121	35.026
	0.1	500	0.068	9.735	0.171	9.674
	0.4	100	0.408	1492.331	0.343	1475.090
	0.4	200	0.436	551.640	0.546	549.310
	0.4	500	0.293	155.513	0.689	154.893

Continued...

Population	σ	n	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
8.Fast sine	0.1	100	0.001	160.854	0.001	155.795
	0.1	200	0.001	42.423	0.001	41.988
	0.1	500	0.001	10.692	0.001	10.654
	0.4	100	0.004	1717.777	0.003	1690.978
	0.4	200	0.004	611.102	0.003	607.835
	0.4	500	0.003	160.394	0.005	159.522

Table 1.2: CV smoothing parameters $\widehat{\alpha}_{\mathrm{CV}}$ and optimal smoothing parameters α_{opt} with their corresponding MSEs based on 1000 replications of simple random sampling from all populations of size $N = 1000$.

Table 1.2 shows the cross-validation smoothing parameters $\widehat{\alpha}_{\mathrm{CV}}$ and the optimal smoothing parameters α_{opt} for linear spline regression under SI simulation runs. Apparently, in agreement with α_{opt}, $\widehat{\alpha}_{\mathrm{CV}}$ varies widely across functions. CV and optimal smoothing parameters are generally an increasing function of the closeness of the relationship between y and x (as measured by σ^2). The difference between CV and optimal smoothing parameters is generally a decreasing function of the sample size.

Population	σ	n	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
1.Linear	0.1	100	8.509	55.090	10.000	54.552
	0.1	200	8.903	24.900	10.000	24.711
	0.1	500	8.614	7.313	10.000	7.234
	0.4	100	8.509	881.432	10.000	872.835
	0.4	200	8.903	398.395	10.000	395.372
	0.4	500	8.614	117.006	10.000	115.737
2.Quadratic	0.1	100	2.392	56.826	2.205	56.353
	0.1	200	2.582	25.234	1.556	25.160
	0.1	500	1.698	7.321	2.477	7.295
	0.4	100	7.590	881.681	10.000	875.107
	0.4	200	8.127	399.683	5.591	397.601
	0.4	500	7.149	115.910	10.000	115.519
3.Bump	0.1	100	0.004	88.472	0.008	82.967
	0.1	200	0.004	33.318	0.007	31.857
	0.1	500	0.003	8.754	0.006	8.556
	0.4	100	0.194	1103.057	0.107	1028.741
	0.4	200	0.044	464.292	0.053	447.779
	0.4	500	0.017	128.998	0.038	126.544

Continued...

Population	σ	n	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
4.Jump	0.1	100	0.012	117.315	0.007	107.400
	0.1	200	0.001	44.271	0.007	38.595
	0.1	500	0.000	10.591	0.003	9.675
	0.4	100	2.303	1032.115	0.774	993.904
	0.4	200	0.936	434.021	0.343	422.015
	0.4	500	0.131	128.821	0.385	120.176
5.Normal CDF	0.1	100	7.351	55.906	10.000	55.024
	0.1	200	7.218	24.797	10.000	24.734
	0.1	500	4.691	7.383	10.000	7.283
	0.4	100	8.388	887.062	10.000	873.790
	0.4	200	8.770	396.790	10.000	395.243
	0.4	500	8.097	117.254	10.000	115.877
6.Exponential	0.1	100	0.287	64.886	0.171	61.646
	0.1	200	0.183	27.055	0.192	26.299
	0.1	500	0.064	7.750	0.242	7.535
	0.4	100	5.422	915.976	3.944	903.500
	0.4	200	4.842	405.317	2.477	400.013
	0.4	500	2.913	116.866	3.944	115.908
7.Slow sine	0.1	100	0.090	63.817	0.107	62.343
	0.1	200	0.089	27.013	0.135	26.468
	0.1	500	0.041	7.812	0.085	7.668
	0.4	100	0.399	962.404	0.433	938.687
	0.4	200	0.390	412.783	0.486	409.177
	0.4	500	0.208	120.576	0.385	119.704
8.Fast sine	0.1	100	0.021	153.564	0.001	138.172
	0.1	200	0.001	42.696	0.002	41.585
	0.1	500	0.001	10.157	0.001	9.924
	0.4	100	0.211	1599.195	0.005	1426.315
	0.4	200	0.005	538.219	0.007	519.818
	0.4	500	0.003	139.213	0.007	135.693

Table 1.3: CV smoothing parameters $\widehat{\alpha}_{\mathrm{CV}}$ and optimal smoothing parameters α_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling from all populations of size $N = 1000$ with $r = 0.5$.

The same overall behavior can be seen in the results of Table 1.3, which displays the CV and optimal smoothing parameters for linear spline regression estimation under the STSI design with $r = 0.5$.

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
1.Linear	0	8.658	41.252	10.000	41.152
	0.25	8.515	34.379	10.000	34.178
	0.5	8.903	24.900	10.000	24.711
	0.75	9.129	17.116	10.000	17.062
	1	9.541	7.150	10.000	7.129
2.Quadratic	0	2.334	41.733	2.205	41.442
	0.25	2.617	35.032	2.477	34.749
	0.5	2.582	25.234	1.556	25.160
	0.75	2.154	18.001	2.477	17.792
	1	1.587	7.858	2.783	7.872
3.Bump	0	0.005	44.122	0.004	43.908
	0.25	0.004	40.491	0.006	40.153
	0.5	0.004	33.318	0.007	31.857
	0.75	0.004	28.286	0.009	26.853
	1	0.004	21.716	0.012	19.160
4.Jump	0	0.002	51.587	0.001	50.571
	0.25	0.001	50.233	0.001	48.406
	0.5	0.001	44.271	0.007	38.595
	0.75	0.001	41.116	0.006	35.384
	1	0.001	38.558	0.009	29.025
5.Normal CDF	0	7.465	41.459	10.000	41.274
	0.25	7.169	34.481	10.000	34.135
	0.5	7.218	24.797	10.000	24.734
	0.75	7.107	17.299	10.000	17.158
	1	7.494	7.341	10.000	7.312
6.Exponential	0	0.182	43.397	0.038	42.844
	0.25	0.204	37.208	0.135	36.496
	0.5	0.183	27.055	0.192	26.299
	0.75	0.152	20.027	0.242	19.801
	1	0.119	10.221	0.305	9.731
7.Slow sine	0	0.093	42.316	0.095	42.383
	0.25	0.085	36.766	0.192	35.954
	0.5	0.089	27.013	0.135	26.468
	0.75	0.087	20.523	0.152	20.093
	1	0.093	10.917	0.171	10.629
8.Fast sine	0	0.001	48.984	0.001	48.127
	0.25	0.001	47.002	0.001	46.631
	0.5	0.001	42.696	0.002	41.585
	0.75	0.001	39.406	0.002	37.901
	1	0.001	37.149	0.002	34.077

Continued. . .

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$

Table 1.4: CV smoothing parameters $\widehat{\alpha}_{\mathrm{CV}}$ and optimal smoothing parameters α_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling from populations of size $N = 1000$ with model variance $\sigma^2 = 0.01$, and the stratum sample sizes $n = (30, 40, 60, 70)$.

Table 1.4 displays the CV smoothing parameters $\widehat{\alpha}_{\mathrm{CV}}$ and optimal smoothing parameters α_{opt} with their corresponding MSEs under the design of stratified simple random sampling without replacement for linear spline regression estimation with model variance $\sigma^2 = 0.01$ and the stratum sample sizes $n = (30, 40, 60, 70)$. We can find that $\widehat{\alpha}_{\mathrm{CV}}$ changes within a short range, and it tracks α_{opt} well. Because the decrease of r means that the relationship between the inclusion probability and the model error becomes weaker, it is reasonable to find that the MSE decreases as r increases. Also, the optimal α and its estimator remain stable across values of r, so the method we propose works even when the design effect is important.

From the above tables, we can see that the smoothing parameter selection method provides a reasonably accurate estimate of the minimizer of the MSE. Even in the cases where it leads to a value further from the true minimizer, it appears to perform well in the sense of selecting a smoothing parameter leading to the best possible MSE. This finding holds across the sample sizes, model variances and sampling design methods we considered. Generally, when the MSE function has a minimum value based on the smoothing parameter, the CV criterion (1.13) performs successfully in approximating the minimum of the MSE.

1.5 Conclusion

In this chapter, we proposed a design-based CV criterion for selecting the smoothing parameter in penalized spline regression estimation. First, we developed theoretical results by proving that the design-based properties of the P-spline regression estimator hold uniformly for a range of smoothing parameter values, making smoothing parameter selection possible.

Then those results were applied to show that the proposed method for smoothing parameter selection is asymptotically equivalent to minimizing the MSE of the estimator. By a simulation study, we showed that the estimated smoothing parameter usually tracks the optimal parameter quite well. Hence, we recommend the design-based CV criterion whenever data-driven smoothing parameter selection for survey estimation is required, except that alternative methods are developed for smoothing parameter selection in the finite population estimation context.

CHAPTER 2. CV as Improved Variance Estimation for Model-Assisted Estimators

2.1 Introduction

A characteristic of sampling survey is to use the available auxiliary information of the population to improve the precision of design-based estimators. The regression estimators have been used for a long time in survey estimation to make auxiliary information be efficiently used.

Särndal (1982) proposed a variance estimator for the regression estimator. And Särndal, Swensson, and Wretman (1989) studied its properties with respect to the design and the model respectively. It has been shown that the estimator is design consistent and approximately unbiased.

In this chapter, we will propose a new variance estimator based on the "leave-one-out" or cross-validation (CV) principle following the construction of CV criterion in Opsomer and Miller (2005). The purpose of this chapter is to introduce the new variance estimator, explore its general properties and compare it to other variance estimators used currently.

Section 2.2 will give the definition of the regression estimator and introduce the CV-based variance estimator. In section 2.3 we state assumptions used in the theoretical derivations and our main theoretical results are described. In section 2.4, we report simulation results, which show how well the CV-based variance estimator works in practice.

2.2 Definition of the Estimator

The estimation of a finite population mean is considered,

$$\bar{y}_N = \frac{1}{N} \sum_{i \in U} y_i.$$

Similar to section 1.1 $U = \{1, 2, ..., N\}$ is a finite population with N identifiable elements and y_i is a response variable for the ith element. A sample of population elements $s \subset U$ is selected with probability $p(s)$. Let $\pi_i = \Pr(i \in s) = \sum_{s: i \in s} p(s) > 0$ denote the inclusion probability for element i. Then the π estimator of $\bar{y}_N$ is

$$\bar{y}_\pi = \frac{1}{N} \sum_{s} \frac{y_i}{\pi_i}, \tag{2.1}$$

which is based on the Horvitz-Thompson estimator (Horvitz and Thompson 1952). The variance of $\bar{y}_\pi$ under sampling design is

$$\mathrm{Var}(\bar{y}_\pi) = \frac{1}{N^2} \sum_{i \in U} \sum_{j \in U} (\pi_{ij} - \pi_i \pi_j) \frac{y_i}{\pi_i} \frac{y_j}{\pi_j}, \tag{2.2}$$

where $\pi_{ij} = \Pr(i \in s, j \in s)$, the joint inclusion probability for elements $i, j \in U$.

Suppose there is auxiliary information x_i available for all of U. And the value of the auxiliary variable vector for the ith element is denoted by

$$\boldsymbol{x}_i = (x_{i1}, \ldots, x_{iJ})^T,$$

where J is the number of auxiliary variable. Then the general regression estimator, denoted by $\bar{y}_{\text{reg}}$, is defined from $\widehat{t}_{yr}$ in Särndal et al. (1992, p. 225) as follows

$$\bar{y}_{\text{reg}} = \bar{y}_\pi + \left(\bar{\boldsymbol{X}}_N - \bar{\boldsymbol{X}}_\pi\right)^T \widehat{\boldsymbol{\beta}}, \tag{2.3}$$

where

$$\bar{\boldsymbol{X}}_\pi = \frac{1}{N} \sum_{s} \frac{\boldsymbol{x}_i}{\pi_i} = \frac{1}{N} \widehat{\boldsymbol{t}}_{x\pi}$$

is the π estimator of the known $\bar{\boldsymbol{X}}_N$,

$$\bar{\boldsymbol{X}}_N = \frac{1}{N} \sum_{U} \boldsymbol{x}_i = \frac{1}{N} \boldsymbol{t}_x.$$

And $\widehat{\boldsymbol{\beta}}$ is

$$\widehat{\boldsymbol{\beta}} = \left(\widehat{\beta}_1, \ldots, \widehat{\beta}_J\right)^T = \left(\sum_s \boldsymbol{x}_i \boldsymbol{x}_i^T / \sigma_i^2 \pi_i\right)^{-1} \sum_s \boldsymbol{x}_i y_i / \sigma_i^2 \pi_i \tag{2.4}$$

with the σ_i^2 assumed known up to a proportionality constant.

As shown in (2.3), the regression estimator is explicitly the π estimator plus an adjustment term. The adjustment term is often negatively correlated with the error of the π estimator when the regression estimator works well.

The regression model ξ motivating (2.3), with y as the regressor and $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_J$ as regressands, will have the following features:

i. $y_1, \ldots, y_N$ are assumed to be realized values of independent random variables $Y_1, \ldots, Y_N$,

ii. $\mathrm{E}_\xi\left(Y_i\right) = \boldsymbol{x}_i^T \boldsymbol{\beta}$,

iii. $V_\xi\left(Y_i\right) = \sigma_i^2$ $(i = 1, \ldots, N)$,

where E_ξ and V_ξ denote expected value and variance with respect to the model ξ, and $\boldsymbol{\beta}$ and σ_i^2 are model parameters. Under the model ξ the population-level weighted least-squares estimator of

$$\boldsymbol{\beta} = \left(\beta_1, \ldots, \beta_J\right)^T$$

will be

$$\boldsymbol{\beta}_N = \left(\beta_{N1}, \ldots, \beta_{NJ}\right)^T = \left(\sum_U \boldsymbol{x}_i \boldsymbol{x}_i^T / \sigma_i^2\right)^{-1} \sum_U \boldsymbol{x}_i y_i / \sigma_i^2, \tag{2.5}$$

which could be written as more familiar expression from regression analysis,

$$\boldsymbol{\beta}_N = \left(\boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{V}^{-1} \boldsymbol{Y},$$

where $\boldsymbol{X}$ represents the matrix of dimension $N \times J$ with rows $\boldsymbol{x}_i^T = (x_{i1}, \ldots, x_{iJ})$, $\boldsymbol{Y} = (y_1, \ldots, y_N)^T$, and $\boldsymbol{V}$ is $N \times N$ diagonal matrix

$$\boldsymbol{V} = \begin{pmatrix} \sigma_1^2 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & \sigma_N^2 \end{pmatrix}.$$

It has been shown that $\boldsymbol{\beta}_N$ is the best linear unbiased estimator of $\boldsymbol{\beta}$ under the model. Note that $\boldsymbol{\beta}_N$ is a finite population characteristic unknown to us, but we can estimate it using sample data by applying "π estimation (inverse-probability-weighting)". Write the unknown $\boldsymbol{\beta}_N$ as

$$\boldsymbol{\beta}_N = \boldsymbol{T}^{-1}\boldsymbol{t}, \tag{2.6}$$

where

$$\boldsymbol{T} = \sum_U \frac{\boldsymbol{x}_i \boldsymbol{x}_i^T}{\sigma_i^2}; \qquad \boldsymbol{t} = \sum_U \frac{\boldsymbol{x}_i y_i}{\sigma_i^2}. \tag{2.7}$$

$\boldsymbol{T}$ is a symmetric $J \times J$ matrix and $\boldsymbol{t}$ is a J-vector. The typical elements of $\boldsymbol{T}$ and $\boldsymbol{t}$ are population product totals denoted respectively by:

$$t_{jj'} = \sum_U \frac{x_{ij} x_{ij'}}{\sigma_i^2} = t_{j'j}; \qquad t_{j0} = \sum_U \frac{x_{ij} y_i}{\sigma_i^2}. \tag{2.8}$$

The π estimators for $\boldsymbol{T}$ and $\boldsymbol{t}$ respectively are

$$\widehat{\boldsymbol{T}} = \sum_s \frac{\boldsymbol{x}_i \boldsymbol{x}_i^T}{\sigma_i^2 \pi_i}; \qquad \widehat{\boldsymbol{t}} = \sum_s \frac{\boldsymbol{x}_i y_i}{\sigma_i^2 \pi_i}. \tag{2.9}$$

Their typical elements are given by:

$$\widehat{t}_{jj'} = \sum_s \frac{x_{ij} x_{ij'}}{\sigma_i^2 \pi_i} = \widehat{t}_{j'j}; \qquad \widehat{t}_{j0} = \sum_s \frac{x_{ij} y_i}{\sigma_i^2 \pi_i}. \tag{2.10}$$

They are unbiased for $t_{jj'}$ and t_{j0} respectively. Then the population parameter $\boldsymbol{\beta}_N$ is estimated by

$$\begin{aligned} \widehat{\boldsymbol{\beta}} &= \widehat{\boldsymbol{T}}^{-1} \widehat{\boldsymbol{t}} \\ &= \left(\sum_s \frac{\boldsymbol{x}_i \boldsymbol{x}_i^T}{\sigma_i^2 \pi_i} \right)^{-1} \sum_s \frac{\boldsymbol{x}_i y_i}{\sigma_i^2 \pi_i}, \end{aligned} \tag{2.11}$$

which is the same expression proposed in (2.4).

Referring to Särndal et al. (1992), note that,

$$\begin{aligned} \bar{y}_{\text{reg}} &= \bar{y}_\pi + \left(\bar{\boldsymbol{X}}_N - \bar{\boldsymbol{X}}_\pi \right)^T \widehat{\boldsymbol{\beta}} \\ &= \frac{1}{N} \sum_{i \in s} g_{is} \frac{y_i}{\pi_i} \\ &= \frac{1}{N} \sum_{i \in U} \widehat{y}_i + \frac{1}{N} \sum_{i \in s} \frac{e_{is}}{\pi_i} \\ &= \frac{1}{N} \sum_{i \in U} y_i^0 + \frac{1}{N} \sum_{i \in s} g_{is} \frac{E_i}{\pi_i}, \end{aligned} \tag{2.12}$$

where $g_{is} = 1 + \left(\boldsymbol{t}_x - \widehat{\boldsymbol{t}}_{x\pi}\right)^T \widehat{\boldsymbol{T}}^{-1} \boldsymbol{x}_i / \sigma_i^2$, $e_{is} = y_i - \widehat{y}_i = y_i - \boldsymbol{x}_i^T \widehat{\boldsymbol{\beta}}$, and $E_i = y_i - y_i^0 = y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}_N$.

The regression estimator $\bar{y}_{\text{reg}}$ is approximated through Taylor linearization by

$$\begin{aligned} \bar{y}_{\text{reg},0} &= \bar{y}_\pi + \left(\bar{\boldsymbol{X}}_N - \bar{\boldsymbol{X}}_\pi\right)^T \boldsymbol{\beta}_N \\ &= \frac{1}{N} \sum_{i \in U} y_i^0 + \frac{1}{N} \sum_{i \in s} \check{E}_i, \end{aligned} \tag{2.13}$$

where $\check{E}_i = E_i / \pi_i$. Then the approximate variance is given by

$$AV\left(\bar{y}_{\text{reg}}\right) = \frac{1}{N^2} \sum_{i,j \in U} \Delta_{ij} \check{E}_i \check{E}_j, \tag{2.14}$$

with $\Delta_{ij} = \pi_{ij} - \pi_i \pi_j$. Replacing the unobservable $\check{E}_i$ by the observable $\check{e}_{is}$, we obtain the "naive" variance estimator

$$\widehat{V}_n\left(\bar{y}_{\text{reg}}\right) = \frac{1}{N^2} \sum_{i,j \in s} \check{\Delta}_{ij} \check{e}_{is} \check{e}_{js}, \tag{2.15}$$

where $\check{e}_{is} = e_{is} / \pi_i$.

From (2.12), it follows that $V\left(\bar{y}_{\text{reg}}\right) = \frac{1}{N^2} V\left(\sum_{i \in s} g_{is} \check{E}_i\right)$. Disregarding that the weights g_{is} are sample dependent, and inserting $\check{e}_{is}$ for $\check{E}_i$, we obtain the "g-corrected" variance estimator:

$$\widehat{V}_g\left(\bar{y}_{\text{reg}}\right) = \frac{1}{N^2} \sum_{i,j \in s} \check{\Delta}_{ij} \left(g_{is} \check{e}_{is}\right) \left(g_{js} \check{e}_{js}\right). \tag{2.16}$$

Both estimators in (2.15) and (2.16) are based on large sample approximations. And for a given level of $1 - \alpha$, either of the two variance estimators gives approximately $100\,(1 - \alpha)\,\%$ coverage rate in repeated large samples. These intervals are also based on asymptotic normality. However, the available evidence in a number of cases suggests that (2.16) is preferable. This estimator was proposed in Särndal (1982). Särndal, Swensson, and Wretman (1989) studied its properties with respect to the design and the model respectively. It has been shown that the estimator (2.16) is design consistent and approximately unbiased.

In this chapter, we propose a new variance estimator based on the "leave-one-out" or cross-validation (CV) principle. The CV variance estimator for $\bar{y}_{\text{reg}}$ is given by

$$\widehat{V}_{\text{CV}}\left(\bar{y}_{\text{reg}}\right) = \frac{1}{N^2} \sum_{i,j \in s} \check{\Delta}_{ij} \check{e}_{is}^{(-)} \check{e}_{js}^{(-)},$$

where $\check{e}_{is}^{(-)} = e_{is}^{(-)}/\pi_i$, $e_{is}^{(-)} = y_i - \widehat{y}_i^{(-)}$. In this expression, we replace $\widehat{y}_i$ by the "leave-one-out" estimator $\widehat{y}_i^{(-)}$. It can be shown that

$$\begin{aligned}
\widehat{y}_i^{(-)} &= \boldsymbol{x}_i^T \left(\widehat{\boldsymbol{T}} - \frac{\boldsymbol{x}_i \boldsymbol{x}_i^T}{\sigma_i^2 \pi_i} \right)^{-1} \left(\widehat{\boldsymbol{t}} - \frac{\boldsymbol{x}_i y_i}{\sigma_i^2 \pi_i} \right) \\
&= \boldsymbol{x}_i^T \left(\widehat{\boldsymbol{T}}^{-1} + \frac{\widehat{\boldsymbol{T}}^{-1} \frac{\boldsymbol{x}_i \boldsymbol{x}_i^T}{\sigma_i^2 \pi_i} \widehat{\boldsymbol{T}}^{-1}}{1 - w_{sii}^*} \right) \left(\widehat{\boldsymbol{t}} - \frac{\boldsymbol{x}_i y_i}{\sigma_i^2 \pi_i} \right) \\
&= \widehat{y}_i - w_{sii}^* y_i + \frac{w_{sii}^* \widehat{y}_i - \left(w_{sii}^*\right)^2 y_i}{1 - w_{sii}^*} \\
&= \frac{\widehat{y}_i - w_{sii}^* y_i}{1 - w_{sii}^*},
\end{aligned} \tag{2.17}$$

where $w_{sii}^* = \boldsymbol{x}_i^T \widehat{\boldsymbol{T}}^{-1} \boldsymbol{x}_i / \sigma_i^2 \pi_i$. Which implies that

$$\begin{aligned}
e_{is}^{(-)} &= y_i - \widehat{y}_i^{(-)} \\
&= \frac{y_i - \widehat{y}_i}{1 - w_{sii}^*} \\
&= \frac{e_{is}}{1 - w_{sii}^*}.
\end{aligned} \tag{2.18}$$

Then, $\widehat{V}_{\mathrm{CV}}\left(\bar{y}_{\mathrm{reg}}\right)$ can be expressed by

$$\widehat{V}_{\mathrm{CV}}\left(\bar{y}_{\mathrm{reg}}\right) = \frac{1}{N^2} \sum_{i,j \in s} \check{\Delta}_{ij} \frac{e_{is}}{\pi_i} \frac{e_{js}}{\pi_j} \frac{1}{1 - w_{sii}^*} \frac{1}{1 - w_{sjj}^*}. \tag{2.19}$$

2.3 Theoretical Properties

In order to prove our theoretical results, we make the following technical assumptions. For simplicity, we will only consider the case with the sample size, denoted by n_N, fixed for each N, and also assume that $n_N \to \infty$. As above, J is fixed.

A1. (Sampling rate $n_N N^{-1}$). As $N \to \infty$, $n_N N^{-1} \to \pi \in (0, 1)$.

A2. (Inclusion probabilities π_i and π_j). For all N, $\min_{i \in U} \pi_i \geq \lambda > 0$, $\min_{i,j \in U} \pi_{ij} \geq \lambda^* > 0$, $\lim_{N \to \infty} n_N \max_{i,j \in U, i \neq j} |\Delta_{ij}| < \infty$.

A3. Assume that $\left(\frac{n_N}{N^2} \widehat{\boldsymbol{T}}\right)^{-1}$ and $\left(\frac{n_N}{N^2} \boldsymbol{T}\right)^{-1}$ exist for all samples.

A4. $\max_{i\in U}|y_i| < C_y$, and $\max_{i\in U, j\in\{1,\ldots,J\}}|x_{ij}| < C_x$, where C_y and C_x are some positive constants.

A5. $0 < \sigma_L \le \min_{i\in U}\sigma_i \le \max_{i\in U}\sigma_i \le \sigma_U < \infty$, where σ_L and σ_U are some positive constants.

A6. Additional assumptions involving higher-order inclusion probabilities:

$$\lim_{N\to\infty}\max_{(i,j,k)\in D_{3,N}}\left|\mathrm{E}_p\left[\left(I_i-\pi_i\right)\left(I_jI_k-\pi_{jk}\right)\right]\right| = 0$$

$$\lim_{N\to\infty}\max_{(i,j,k,l)\in D_{4,N}}\left|\mathrm{E}_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_kI_l-\pi_{kl}\right)\right]\right| = 0$$

$$\lim_{N\to\infty} n_N \max_{(i,j,k)\in D_{3,N}}\left|\mathrm{E}_p\left[\left(I_i-\pi_i\right)^2\left(I_j-\pi_j\right)\left(I_k-\pi_k\right)\right]\right| < \infty$$

$$\lim_{N\to\infty} n_N^2 \max_{(i,j,k,l)\in D_{4,N}}\left|\mathrm{E}_p\left[\left(I_i-\pi_i\right)\left(I_j-\pi_j\right)\left(I_k-\pi_k\right)\left(I_l-\pi_l\right)\right]\right| < \infty$$

where $D_{t,N}$ denotes the set of all distinct t-tuples from U.

Assumption A3 ensures that $\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{\beta}_N$ exist. The following results establish design consistency of variance estimator of $\bar{y}_{\mathrm{reg}}$.

Theorem 2.3.1. *Let assumptions A1-A6 hold. Then, we have the following result:*

$$\begin{aligned}\mathrm{MSE}_p\left(\bar{y}_{\mathrm{reg}}\right) &= \mathrm{E}_p\left(\bar{y}_{\mathrm{reg}}-\bar{y}_N\right)^2 \\ &= \frac{1}{N^2}\sum_{i,j\in U}\Delta_{ij}\frac{E_i}{\pi_i}\frac{E_j}{\pi_j} + o\left(n_N^{-1}\right).\end{aligned}$$

Proof of Theorem 2.3.1: Let $a_N = \sum_U\left(y_i - y_i^0\right)\left(\frac{I_i}{\pi_i}-1\right)$, $b_N = \sum_U\left(y_i^0 - \widehat{y}_i\right)\left(\frac{I_i}{\pi_i}-1\right)$. Then,

$$\begin{aligned}\mathrm{E}_p\left[a_N^2\right] &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{y_j-y_j^0}{\pi_j} \\ &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{E_i}{\pi_i}\frac{E_j}{\pi_j}.\end{aligned}$$

And by assumptions A1, A2 and Lemma A.2.3,

$$\begin{aligned}
\mathrm{E}_p\left[a_N^2\right] &\leq \frac{1}{\lambda^2}\sum_{i\in U}\max_{i\in U}\left|\pi_i\left(1-\pi_i\right)\right|\left\{\max_{i\in U}\left|y_i-y_i^0\right|\right\}^2 \\
&\quad +\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\left\{\max_{i\in U}\left|y_i-y_i^0\right|\right\}^2 \\
&= O\left(N\right)+O\left(\frac{N^2}{n_N}\right) \\
&= O\left(\frac{N^2}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U}\left|\Delta_{ii}\right| = \max_{i\in U}\left|\pi_i\left(1-\pi_i\right)\right| \leq 1$. Then it follows that

$$\frac{1}{N^2}\mathrm{E}_p\left[a_N^2\right] = O\left(\frac{1}{n_N}\right).$$

Let b_k denote the kth element of $\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_N$, from assumptions A1, A2, A4, A6, by Lemma A.2.2 we can show that

$$\begin{aligned}
\mathrm{E}_p\left[b_N^2\right] &= \mathrm{E}_p\left[\sum_{(i,j)\in U}\sum_{k,l\in\{1,\ldots,J\}} x_{ik}x_{jl}b_kb_l\left(\frac{I_i}{\pi_i}-1\right)\left(\frac{I_j}{\pi_j}-1\right)\right] \\
&\le \sum_{k,l\in\{1,\ldots,J\}}\sqrt{\mathrm{E}_p\left[b_k^2b_l^2\right]}\sqrt{\mathrm{E}_p\left[\left\{\sum_{(i,j)\in U} x_{ik}x_{jl}\left(\frac{I_i}{\pi_i}-1\right)\left(\frac{I_j}{\pi_j}-1\right)\right\}^2\right]} \\
&\le \sum_{k,l\in\{1,\ldots,J\}}\frac{1}{\lambda^2}\sqrt{\max_{k\in\{1,\ldots,J\}} E_p\left[b_k^4\right]}\left\{\sqrt{\sum_{i\in U}\left\{\max_{i\in U,k\in\{1,\ldots,J\}}|x_{ik}|\right\}^4\max_{i\in U}\left|\mathrm{E}_p\left[(I_i-\pi_i)^4\right]\right|}\right. \\
&\quad +\sqrt{\sum_{(i,j)\in D_{2,N}}\left\{\max_{i\in U,k\in\{1,\ldots,J\}}|x_{ik}|\right\}^4\max_{(i,j)\in D_{2,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^2(I_j-\pi_j)^2\right]\right|} \\
&\quad +\sqrt{\sum_{(i,j)\in D_{2,N}}\left\{\max_{i\in U,k\in\{1,\ldots,J\}}|x_{ik}|\right\}^4\max_{(i,j)\in D_{2,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^3(I_j-\pi_j)\right]\right|} \\
&\quad +\sqrt{\sum_{(i,j,i')\in D_{3,N}}\left\{\max_{i\in U,k\in\{1,\ldots,J\}}|x_{ik}|\right\}^4\max_{(i,j,i')\in D_{3,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^2(I_j-\pi_j)(I_{i'}-\pi_{i'})\right]\right|} \\
&\quad +\sqrt{\sum_{(i,j,i',j')\in D_{4,N}}\left\{\max_{i\in U,k\in\{1,\ldots,J\}}|x_{ik}|\right\}^4} \\
&\quad \left.\times\sqrt{\max_{(i,j,i',j')\in D_{4,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)(I_j-\pi_j)(I_{i'}-\pi_{i'})(I_{j'}-\pi_{j'})\right]\right|}\right\} \\
&= O\left(\frac{1}{n_N}\right)O\left(\sqrt{N}\right)+O\left(\frac{1}{n_N}\right)O(N)+O\left(\frac{1}{n_N}\right)O(N)+O\left(\frac{1}{n_N}\right)O\left(\frac{N^{3/2}}{\sqrt{n_N}}\right) \\
&\quad +O\left(\frac{1}{n_N}\right)O\left(\frac{N^2}{n_N}\right) \\
&= o\left(\frac{N^2}{n_N}\right),
\end{aligned}$$

which implies that

$$\frac{1}{N^2}\mathrm{E}_p\left[b_N^2\right] = o\left(\frac{1}{n_N}\right).$$

and,

$$\begin{aligned}
\frac{1}{N^2}\mathrm{E}_p\left[|a_Nb_N|\right] &\le \sqrt{\frac{1}{N^2}\mathrm{E}_p\left[a_N^2\right]\frac{1}{N^2}\mathrm{E}_p\left[b_N^2\right]} \\
&= \sqrt{O\left(\frac{1}{n_N}\right)o\left(\frac{1}{n_N}\right)} \\
&= o\left(\frac{1}{n_N}\right).
\end{aligned}$$

Rewrite $\mathrm{MSE}_p(\bar{y}_{\mathrm{reg}})$ as follows,

$$\begin{aligned}
\mathrm{MSE}_p(\bar{y}_{\mathrm{reg}}) &= \mathrm{E}_p\left[(\bar{y}_{\mathrm{reg}} - \bar{y}_N)^2\right] \\
&= \frac{1}{N^2}\mathrm{E}_p\left[\left(\sum_{i\in U}\widehat{y}_i + \sum_{i\in s}\frac{y_i - \widehat{y}_i}{\pi_i} - \sum_{i\in U} y_i\right)^2\right] \\
&= \frac{1}{N^2}\mathrm{E}_p\left[\left(\sum_{i\in U}(y_i - \widehat{y}_i)\left(\frac{I_i}{\pi_i} - 1\right)\right)^2\right] \\
&= \frac{1}{N^2}\mathrm{E}_p\left[\left(\sum_{i\in U}\left(y_i - y_i^0 + y_i^0 - \widehat{y}_i\right)\left(\frac{I_i}{\pi_i} - 1\right)\right)^2\right] \\
&= \frac{1}{N^2}\mathrm{E}_p\left[(a_N + b_N)^2\right] \\
&= \frac{1}{N^2}\left\{\mathrm{E}_p\left[a_N^2\right] + \mathrm{E}_p\left[b_N^2\right] + 2\mathrm{E}_p\left[a_N b_N\right]\right\}.
\end{aligned}$$

Then,

$$\begin{aligned}
\mathrm{MSE}_p(\bar{y}_{\mathrm{reg}}) &= \frac{1}{N^2}\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{E_i}{\pi_i}\frac{E_j}{\pi_j} + o\left(\frac{1}{n_N}\right) + o\left(\frac{1}{n_N}\right) \\
&= \frac{1}{N^2}\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{E_i}{\pi_i}\frac{E_j}{\pi_j} + o\left(\frac{1}{n_N}\right).
\end{aligned}$$

Therefore the result follows.

□

Theorem 2.3.2. *Under assumptions A1-A6, we have that the estimator*

$$\begin{aligned}
\widehat{V}_{\mathrm{CV}}(\bar{y}_{\mathrm{reg}}) &= \frac{1}{N^2}\sum_{i,j\in s}\check{\Delta}_{ij}\frac{e_{is}}{\pi_i}\frac{e_{js}}{\pi_j}\frac{1}{1 - w_{sii}^*}\frac{1}{1 - w_{sjj}^*} \\
&= \frac{1}{N^2}\sum_{i,j\in U}\Delta_{ij}\frac{E_i}{\pi_i}\frac{E_j}{\pi_j} + o_p\left(n_N^{-1}\right).
\end{aligned}$$

Proof of Theorem 2.3.2: We write the expression as follows

$$
\begin{aligned}
\widehat{V}_{\mathrm{CV}}\left(\bar{y}_{\mathrm{reg}}\right) &= \frac{1}{N^2}\sum_{i\in s}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\frac{y_i-\widehat{y}_i}{\pi_i}\frac{y_j-\widehat{y}_j}{\pi_j}\frac{1}{1-w^*_{sii}}\frac{1}{1-w^*_{sjj}} \\
&= \frac{1}{N^2}\sum_{i\in s}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\frac{y_i-y_i^0}{\pi_i}\frac{y_j-y_j^0}{\pi_j}\frac{1}{1-w^*_{sii}}\frac{1}{1-w^*_{sjj}} \\
&\quad+\frac{1}{N^2}\sum_{i\in s}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\frac{y_i^0-\widehat{y}_i}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j}\frac{1}{1-w^*_{sii}}\frac{1}{1-w^*_{sjj}} \\
&\quad+\frac{2}{N^2}\sum_{i\in s}\sum_{j\in s}\frac{\Delta_{ij}}{\pi_{ij}}\frac{y_i-y_i^0}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j}\frac{1}{1-w^*_{sii}}\frac{1}{1-w^*_{sjj}} \\
&= \frac{1}{N^2}\left(V_1+V_2+2V_3\right).
\end{aligned}
$$

In this expression,

$$
\begin{aligned}
V_1 &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{y_j-y_j^0}{\pi_j}\frac{1}{1-w^*_{sii}}\frac{1}{1-w^*_{sjj}} \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{y_j-y_j^0}{\pi_j}\frac{1}{1-w^*_{sii}}\frac{1}{1-w^*_{sjj}}\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\
&= V_{11}+V_{12}.
\end{aligned}
$$

And,

$$
\begin{aligned}
V_{11} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{y_j-y_j^0}{\pi_j} \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{y_j-y_j^0}{\pi_j}\left(\frac{1}{1-w^*_{sii}}\frac{1}{1-w^*_{sjj}}-1\right) \\
&= \sum_{i,j\in U}\Delta_{ij}\frac{E_i}{\pi_i}\frac{E_j}{\pi_j}+V^*_{11}.
\end{aligned}
$$

Suppose assumptions A1, A2 hold, from Lemma A.2.3 and A.2.6, we can show that

$$
\begin{aligned}
|V_{11}^*| &\leq \sum_{i\in U} |\pi_i (1-\pi_i)| \frac{(y_i - y_i^0)^2}{\pi_i^2} \left| g_{ii}^* \left(\widehat{\boldsymbol{T}}\right)\right| + \\
&\quad \sum_{(i,j)\in D_{2,N}} |\Delta_{ij}| \left| \frac{y_i - y_i^0}{\pi_i} \frac{y_j - y_j^0}{\pi_j} \right| \left| g_{ij}^* \left(\widehat{\boldsymbol{T}}\right)\right| \\
&\leq \frac{1}{\lambda^2} \sum_{i\in U} \left\{ \max_{i\in U} \left| y_i - y_i^0 \right| \right\}^2 \max_{i,j\in U} \left| g_{ij}^* \left(\widehat{\boldsymbol{T}}\right)\right| \\
&\quad + \frac{1}{\lambda^2} \sum_{(i,j)\in D_{2,N}} \max_{(i,j)\in D_{2,N}} |\Delta_{ij}| \left\{ \max_{i\in U} \left| y_i - y_i^0 \right| \right\}^2 \max_{i,j\in U} \left| g_{ij}^* \left(\widehat{\boldsymbol{T}}\right)\right| \\
&= O(1) + O\left(\frac{N}{n_N}\right) \\
&= O\left(\frac{N}{n_N}\right),
\end{aligned}
$$

as the fact that $\max_{i\in U} |\Delta_{ii}| = \max_{i\in U} |\pi_i (1-\pi_i)| \leq 1$. Rewrite V_{12} as follows,

$$
\begin{aligned}
V_{12} &= \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{y_i - y_i^0}{\pi_i} \frac{y_j - y_j^0}{\pi_j} \left(\frac{I_i I_j}{\pi_{ij}} - 1 \right) \\
&\quad + \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{y_i - y_i^0}{\pi_i} \frac{y_j - y_j^0}{\pi_j} \left(\frac{1}{1-w_{sii}^*} \frac{1}{1-w_{sjj}^*} - 1 \right) \left(\frac{I_i I_j}{\pi_{ij}} - 1 \right) \\
&= V_{121} + V_{122}.
\end{aligned}
$$

In this expression, $\mathrm{E}_p [V_{121}] = 0$ and

$$
\begin{aligned}
\mathrm{Var}_p [V_{121}] &= E_p \left[V_{121}^2 \right] \\
&= \sum_{i,j,k,l\in U} \Delta_{ij}\Delta_{kl} \frac{y_i - y_i^0}{\pi_i} \frac{y_j - y_j^0}{\pi_j} \frac{y_k - y_k^0}{\pi_k} \frac{y_l - y_l^0}{\pi_l} \mathrm{E}_p \left[\left(\frac{I_i I_j}{\pi_{ij}} - 1 \right) \left(\frac{I_k I_l}{\pi_{kl}} - 1 \right) \right].
\end{aligned}
$$

Then, by assumptions A1, A2, A6 and Lemma A.2.3,

$$\begin{aligned}
|\mathrm{Var}_p\left[V_{121}\right]| &\leq \frac{1}{\lambda^4}\sum_{i,j,k,l\in U}\max_{i,j\in U}|\Delta_{ij}|\max_{k,l\in U}|\Delta_{kl}|\left\{\max_{i\in U}\left|y_i-y_i^0\right|\right\}^4 \\
&\quad\times\max_{i,j,k,l\in U}\left|E_p\left[\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\frac{I_kI_l}{\pi_{kl}}-1\right)\right]\right| \\
&= \frac{2}{\lambda^5\lambda^*}\sum_{(i,j,k)\in D_{3,N}}O\left(\frac{1}{n_N}\right)\max_{(i,j,k)\in D_{3,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_k-\pi_k\right)\right]\right| \\
&\quad+\frac{4}{\lambda^4{\lambda^*}^2}\sum_{(i,j,k)\in D_{3,N}}O\left(\frac{1}{n_N^2}\right)\max_{(i,j,k)\in D_{3,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_iI_k-\pi_{ik}\right)\right]\right| \\
&\quad+\frac{4}{\lambda^5\lambda^*}\sum_{(i,j)\in D_{2,N}}O\left(\frac{1}{n_N}\right)\max_{(i,j)\in D_{2,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_i-\pi_i\right)\right]\right| \\
&\quad+\frac{1}{\lambda^6}\sum_{i\in U}O\left(1\right)\max_{i\in U}\left|E_p\left[\left(I_i-\pi_i\right)^2\right]\right| \\
&\quad+\frac{1}{\lambda^4{\lambda^*}^2}\sum_{(i,j,k,l)\in D_{4,N}}O\left(\frac{1}{n_N^2}\right)\max_{(i,j,k,l)\in D_{4,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_kI_l-\pi_{kl}\right)\right]\right| \\
&= o\left(\frac{N^3}{n_N}\right)+O\left(\frac{N^3}{n_N^2}\right)+O\left(\frac{N^2}{n_N}\right)+O\left(N\right)+o\left(\frac{N^4}{n_N^2}\right) \\
&= o\left(\frac{N^4}{n_N^2}\right),
\end{aligned}$$

which implies that $|V_{121}|=o_p\left(\frac{N^2}{n_N}\right)$.

By assumption A2,

$$\begin{aligned}
\max_{i\in U}\left|\frac{I_i}{\pi_i}-1\right| &= \max_{i\in U}\left(1,\frac{1}{\pi_i}-1\right) \\
&\leq \frac{1}{\lambda},
\end{aligned} \tag{2.20}$$

and,

$$\begin{aligned}
\max_{(i,j)\in D_{2,N}}\left|\frac{I_iI_j}{\pi_{ij}}-1\right| &= \max_{i\in U}\left(1,\frac{1}{\pi_{ij}}-1\right) \\
&\leq \frac{1}{\lambda^*}.
\end{aligned} \tag{2.21}$$

Then, V_{122} should have the same order as V_{11}^*. Therefore, it follows that

$$\begin{aligned}
|V_{12}| &\leq |V_{121}| + |V_{122}| \\
&= o_p\left(\frac{N^2}{n_N}\right) + O\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}$$

Next, we will show that

$$|V_2| = o_p\left(\frac{N^2}{n_N}\right).$$

Note

$$\begin{aligned}
V_2 &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i^0-\widehat{y}_i}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*} \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i^0-\widehat{y}_i}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\
&= V_{21}+V_{22}.
\end{aligned}$$

And,

$$\begin{aligned}
V_{21} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i^0-\widehat{y}_i}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j} \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i^0-\widehat{y}_i}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j}\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}-1\right) \\
&= \left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right) \\
&\quad+\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right) \\
&= V_{211}+V_{212}
\end{aligned}$$

Suppose assumptions A1, A2 and A4 hold, by Lemma A.2.1, it can be shown that

$$\begin{aligned}
|V_{211}| &\leq \frac{1}{\lambda^2}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right|^T\sum_{i\in U}\left|\pi_i\left(1-\pi_i\right)\boldsymbol{x}_i\boldsymbol{x}_i^T\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&\quad+\frac{1}{\lambda^2}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right|^T\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}|\Delta_{ij}|\left|\boldsymbol{x}_i\boldsymbol{x}_j^T\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&= o_p(1)\,O(N)\,o_p(1)+o_p(1)\,O\left(\frac{N^2}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N^2}{n_N}\right),
\end{aligned}$$

and by Lemma A.2.6

$$\begin{aligned}
|V_{212}| &\leq \frac{1}{\lambda^2}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right|^T\sum_{i\in U}\left|\pi_i\left(1-\pi_i\right)\boldsymbol{x}_i\boldsymbol{x}_i^T\right|\max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&\quad+\frac{1}{\lambda^2}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right|^T\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\left|\boldsymbol{x}_i\boldsymbol{x}_j^T\right| \\
&\quad\times\max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&= o_p(1)\,O(1)\,o_p(1)+o_p(1)\,O\left(\frac{N}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U}\left|\Delta_{ii}\right|=\max_{i\in U}\left|\pi_i\left(1-\pi_i\right)\right|\leq 1$. Then, it follows that

$$\begin{aligned}
|V_{21}| &\leq |V_{211}|+|V_{212}| \\
&= o_p\left(\frac{N^2}{n_N}\right)+o_p\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}$$

Rewrite V_{22} as

$$\begin{aligned}
V_{22} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i^0-\widehat{y}_i}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j}\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i^0-\widehat{y}_i}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j}\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}-1\right)\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\
&= \left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right) \\
&\quad+\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right) \\
&= V_{221}+V_{222}.
\end{aligned}$$

By (2.20) and (2.21), V_{22} should have the same order as V_{21}. It follows that

$$|V_{22}|=o_p\left(\frac{N^2}{n_N}\right).$$

Therefore, it can be shown that

$$\begin{aligned}|V_2| &\leq |V_{21}| + |V_{22}| \\ &= o_p\left(\frac{N^2}{n_N}\right) + o_p\left(\frac{N^2}{n_N}\right) \\ &= o_p\left(\frac{N^2}{n_N}\right).\end{aligned}$$

Finally, we will show that

$$|V_3| = o_p\left(\frac{N^2}{n_N}\right).$$

Note

$$\begin{aligned}V_3 &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i - y_i^0}{\pi_i}\frac{y_j^0 - \widehat{y}_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*} \\ &\quad + \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i - y_i^0}{\pi_i}\frac{y_j^0 - \widehat{y}_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}\left(\frac{I_iI_j}{\pi_{ij}} - 1\right) \\ &= V_{31} + V_{32},\end{aligned}$$

where

$$\begin{aligned}V_{31} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i - y_i^0}{\pi_i}\frac{y_j^0 - \widehat{y}_j}{\pi_j} \\ &\quad + \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i - y_i^0}{\pi_i}\frac{y_j^0 - \widehat{y}_j}{\pi_j}\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*} - 1\right) \\ &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i - y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}\right) \\ &\quad + \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i - y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}}\right) \\ &= V_{311} + V_{312}.\end{aligned}$$

Suppose assumptions A1, A2 and A4 hold, from Lemma A.2.1 and A.2.3 we can show that

$$\begin{aligned}
|V_{311}| &= \left|\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right)\right| \\
&\le \frac{1}{\lambda^2}\sum_{i\in U}\max_{i\in U}\left|y_i-y_i^0\right|\max_{i\in U}|\boldsymbol{x}_i|^T\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&\quad+\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}|\Delta_{ij}|\max_{i\in U}\left|y_i-y_i^0\right|\max_{j\in U}|\boldsymbol{x}_j|^T\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&= O(N)\,o_p(1)+O\left(\frac{N^2}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N^2}{n_N}\right),
\end{aligned}$$

and,

$$\begin{aligned}
|V_{312}| &= \left|\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right)\right| \\
&\le \frac{1}{\lambda^2}\sum_{i\in U}\max_{i\in U}\left|y_i-y_i^0\right|\max_{i\in U}|\boldsymbol{x}_i|^T\max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&\quad+\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}|\Delta_{ij}|\max_{i\in U}\left|y_i-y_i^0\right|\max_{j\in U}|\boldsymbol{x}_j|^T \\
&\quad\times\max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&= O(1)\,o_p(1)+O\left(\frac{N}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U}|\Delta_{ii}|=\max_{i\in U}|\pi_i(1-\pi_i)|\le 1$. Thereby,

$$\begin{aligned}
|V_{31}| &\le |V_{311}|+|V_{312}| \\
&= o_p\left(\frac{N^2}{n_N}\right)+o_p\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}$$

Write V_{32} as follows

$$\begin{aligned} V_{32} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j}\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\ &\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{y_j^0-\widehat{y}_j}{\pi_j}\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}-1\right)\left(\frac{I_iI_j}{\pi_{ij}}-1\right) \\ &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right) \\ &\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right) \\ &= V_{321}+V_{322}. \end{aligned}$$

Similarly, from (2.20) and (2.21), V_{32} should have the same order as V_{31}. It follows that

$$\begin{aligned} |V_3| &\leq |V_{31}|+|V_{32}| \\ &= o_p\left(\frac{N^2}{n_N}\right)+o_p\left(\frac{N^2}{n_N}\right) \\ &= o_p\left(\frac{N^2}{n_N}\right). \end{aligned}$$

Therefore,

$$\begin{aligned} \widehat{V}_{\text{CV}}\left(\bar{y}_{\text{reg}}\right) &= \frac{1}{N^2}\left\{\sum_{i,j\in U}\Delta_{ij}\frac{E_i}{\pi_i}\frac{E_j}{\pi_j}+O\left(\frac{N}{n_N}\right)+o_p\left(\frac{N^2}{n_N}\right)\right\} \\ &\quad +\frac{1}{N^2}\left\{o_p\left(\frac{N^2}{n_N}\right)+2o_p\left(\frac{N^2}{n_N}\right)\right\} \\ &= \frac{1}{N^2}\sum_{i,j\in U}\Delta_{ij}\frac{E_i}{\pi_i}\frac{E_j}{\pi_j}+o_p\left(\frac{1}{n_N}\right). \end{aligned}$$

Thus, the result follows.

$\square$

Corollary 2.3.1. *Let assumptions A1-A6 hold. Then,*

$$\widehat{V}_{\text{CV}}\left(\bar{y}_{\text{reg}}\right)=\text{MSE}_p\left(\bar{y}_{\text{reg}}\right)+o_p\left(n_N^{-1}\right).$$

Thereby the theory derived above for the regression estimator shows that it is possible to use $\widehat{V}_{\text{CV}}\left(\bar{y}_{\text{reg}}\right)$ as an asymptotically equivalent criterion to $\text{MSE}_p\left(\bar{y}_{\text{reg}}\right)$. In section 2.4, we will evaluate how well this variance estimator works.

2.4 Simulation Results

A random population of $N = 1000$ values of error ε is drawn from $\mathcal{N}(0, 1)$. This one error population is used for all simulations, up to multiplication by σ. Eleven populations of y are generated as follows:

$$y_{il} = m_l(x_i) + \varepsilon_i \qquad 1 \leq i \leq 1000, \qquad 1 \leq l \leq 11,$$

where $\{m_l\}_{l=1}^{11}$ are predefined functions given in the Table 2.1. In mean functions m_1 and m_4, a random population of $N = 1000$ values of x is generated from a uniform distribution on $[0, 1]$. In mean function m_2, a 1000 by 3 matrix of x is generated from a uniform distribution on $[0, 1]$ with independent column vectors. In mean function m_3, a 1000 by 8 matrix of x is generated from a uniform distribution on $[0, 1]$ with independent column vectors. Mean functions m_5 to m_9 are similar to m_2 except that the three column vectors are correlated with correlation specified in Table 2.1. To generate a correlated uniform distribution, first we simulate a population of N by J values of matrix from a multivariate normal distribution with specified correlation structure, and then we calculate the cumulative distribution function (cdf) value for each simulated value based on a univariate normal distribution. These generated values follow a multivariate uniform distribution. The correlation structures specified in Table 2.1 for population functions m_5 to m_9 are used in underlying multivariate normal distribution, not those of multivariate uniform distribution. The main goal is to investigate no correlation, low, medium and high, not the exact numbers. In Mean function m_{10} and m_{11}, the population is defined as a constant β_g. Then $\boldsymbol{x}$ is a vector with gth entry equal to 1 and 0 otherwise, for observation i in group g. For simplicity, both mean functions m_{10} and m_{11} have equal group sizes.

The finite population quantities of interest are $\bar{y}_N = \frac{1}{N}\sum_{i=1}^{1000} y_{il}$ for each l. A linear regression model is conducted for all simulation runs.

The samples are drawn by one of two designs, simple random sampling without replacement (SI) or stratified simple random sampling without replacement (STSI). For each simulation run, $M = 10000$ samples are drawn from $\{(x_i, y_i)\}$. For each sample, we compute the estimator $\bar{y}_{\text{reg}}$, $\widehat{V}_n$, $\widehat{V}_g$, and $\widehat{V}_{\text{CV}}$. A simulation run is determined by sample size n, error

Population	Expression
1.Linear1	$m_1(x) = 2x$
2.Linear2	$m_2(\boldsymbol{x}) = 1 + (.5, 1, 2)\,\boldsymbol{x}$
3.Linear3	$m_3(\boldsymbol{x}) = 1 + (.5, 1, 2, 1.5, 2.5, 1.2, 2.2, 1)\,\boldsymbol{x}$
4.Quadratic	$m_4(x) = 1 + 2(x - 0.5)^2$
5.Collinear1	$m_5(\boldsymbol{x}) = 1 + (.5, 1, 2)\,\boldsymbol{x}$, $\rho = 0.1$ for underlying normal distribution
6.Collinear2	$m_6(\boldsymbol{x}) = 1 + (.5, 1, 2)\,\boldsymbol{x}$, $\rho = 0.3$ for underlying normal distribution
7.Collinear3	$m_7(\boldsymbol{x}) = 1 + (.5, 1, 2)\,\boldsymbol{x}$, $\rho = 0.5$ for underlying normal distribution
8.Collinear4	$m_8(\boldsymbol{x}) = 1 + (.5, 1, 2)\,\boldsymbol{x}$, $\rho = 0.7$ for underlying normal distribution
9.Collinear5	$m_9(\boldsymbol{x}) = 1 + (.5, 1, 2)\,\boldsymbol{x}$, $\rho = 0.98$ for underlying normal distribution
10.Post-stratification1	$m_{10}(\boldsymbol{x}) = (-2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5)\,\boldsymbol{x}$
11.Post-stratification2	$m_{11}(\boldsymbol{x}) = (-2, -1.5, -1, -0.5, 0, 0.5, 1, 1.5, 2, 2.5)\,\boldsymbol{x}$

Table 2.1 Population mean functions.

variance σ^2. For the design of simple random sampling without replacement, simulations are done for $n \in \{100, 200, 500\}$, $\sigma^2 \in \{0.01, 0.16\}$. The design of stratified simple random sampling without replacement uses 4 strata with each stratum containing 250 elements, and the stratification of the strata is based on a random variable z_i and ratio r. First, we generate v_i from a standard normal distribution $\mathcal{N}(0, \sigma_v^2)$ with σ_v^2 satisfying $r = \frac{\sigma^2}{\sigma^2+\sigma_v^2}$. Then z_i's are derived as follows:

$$z_i = \begin{cases} v_i + \varepsilon_i & \text{if } 0 < r < 1 \\ v_i & \text{if } r = 0, \text{ where } \sigma_v^2 = 1 \\ \varepsilon_i & \text{if } r = 1 \end{cases}.$$

After sorting by z_i $(i = 1, \ldots, N)$, the population is separated into 4 strata with boundaries given by equally-spaced quantiles of z. Then, simulations are conducted with the stratum sample sizes $\{(15, 20, 30, 35), (30, 40, 60, 70), (75, 100, 150, 175)\}$, $r \in \{0, 0.25, 0.5, 0.75, 1\}$, $\sigma^2 \in \{0.01, 0.16\}$. Thus, the strata have different sampling rates with the inclusion probability correlated with the model error.

Population	σ	n	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
1.Linear1	0.1	100	1.022	1.027	-2.628	-1.777	1.429
		200	1.021	1.058	0.612	0.993	2.674
		500	1.003	1.018	-0.148	-0.061	0.662
	0.4	100	1.022	1.027	-2.628	-1.777	1.429
		200	1.021	1.058	0.612	0.993	2.674
		500	1.003	1.018	-0.148	-0.061	0.662
2.Linear2	0.1	100	1.045	1.044	-4.761	-2.135	3.327
		200	1.020	1.027	-2.238	-1.059	1.779
		500	0.998	0.975	-1.295	-0.989	0.296
	0.4	100	1.045	1.044	-4.761	-2.135	3.327
		200	1.020	1.027	-2.238	-1.059	1.779
		500	0.998	0.975	-1.295	-0.989	0.296

Continued...

Population	σ	n	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
3.Linear3	0.1	100	0.851	0.641	-17.318	-10.722	-0.177
		200	0.881	0.690	-9.502	-6.413	-0.810
		500	0.945	0.832	-3.031	-2.250	0.535
	0.4	100	0.851	0.641	-17.318	-10.722	-0.177
		200	0.881	0.690	-9.502	-6.413	-0.810
		500	0.945	0.832	-3.031	-2.250	0.535
4.Quadratic	0.1	100	0.997	0.957	-5.169	-2.583	-0.386
		200	0.986	0.952	-3.136	-1.978	-0.753
		500	0.978	0.935	-1.803	-1.509	-0.851
	0.4	100	1.022	1.016	-3.310	-2.106	0.857
		200	1.019	1.038	-0.361	0.180	1.748
		500	1.002	1.006	-0.480	-0.348	0.353
5.Collinear1	0.1	100	1.016	0.976	-6.305	-3.742	1.630
		200	1.005	0.972	-3.349	-2.161	0.614
		500	0.987	0.920	-2.106	-1.798	-0.531
	0.4	100	1.016	0.976	-6.305	-3.742	1.630
		200	1.005	0.972	-3.349	-2.161	0.614
		500	0.987	0.920	-2.106	-1.798	-0.531
6.Collinear2	0.1	100	1.015	0.974	-6.332	-3.759	1.582
		200	1.006	0.971	-3.371	-2.180	0.579
		500	0.987	0.921	-2.094	-1.784	-0.523
	0.4	100	1.015	0.974	-6.332	-3.759	1.582
		200	1.006	0.971	-3.371	-2.180	0.579
		500	0.987	0.921	-2.094	-1.784	-0.523
7.Collinear3	0.1	100	1.014	0.972	-6.361	-3.778	1.537
		200	1.005	0.969	-3.395	-2.202	0.546
		500	0.986	0.921	-2.081	-1.770	-0.514
	0.4	100	1.014	0.972	-6.361	-3.778	1.537
		200	1.005	0.969	-3.395	-2.202	0.546
		500	0.986	0.921	-2.081	-1.770	-0.514
8.Collinear4	0.1	100	1.013	0.970	-6.397	-3.803	1.491
		200	1.005	0.968	-3.424	-2.229	0.510
		500	0.986	0.922	-2.067	-1.756	-0.503
	0.4	100	1.013	0.970	-6.397	-3.803	1.491
		200	1.005	0.968	-3.424	-2.229	0.510
		500	0.986	0.922	-2.067	-1.756	-0.503
9.Collinear5	0.1	100	1.011	0.967	-6.474	-3.859	1.413
		200	1.004	0.964	-3.501	-2.301	0.425
		500	0.986	0.923	-2.050	-1.735	-0.489
	0.4	100	1.011	0.967	-6.474	-3.859	1.413
		200	1.004	0.964	-3.501	-2.301	0.425
		500	0.986	0.923	-2.050	-1.735	-0.489

Continued. . .

Population	σ	n	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
10.Post-ST1	0.1	100	0.993	0.787	-14.346	-9.387	1.256
		200	0.996	0.918	-6.148	-3.521	1.843
		500	0.988	0.985	-1.969	-1.304	1.245
	0.4	100	0.993	0.787	-14.346	-9.387	1.256
		200	0.996	0.918	-6.148	-3.521	1.843
		500	0.988	0.985	-1.969	-1.304	1.245
11.Post-ST2	0.1	100	0.952	0.676	-18.703	-13.135	0.492
		200	0.955	0.826	-8.331	-5.119	1.586
		500	0.962	0.909	-2.906	-2.051	1.098
	0.4	100	0.952	0.676	-18.703	-13.135	0.492
		200	0.955	0.826	-8.331	-5.119	1.586
		500	0.962	0.909	-2.906	-2.051	1.098

Table 2.2: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of simple random sampling from populations of size $N = 1000$.

Table 2.2 shows the ratios of RMSEs δ_g and δ_{CV}, where $\delta_g = \text{RMSE}_g/\text{RMSE}_n$, $\delta_{\text{CV}} = \text{RMSE}_{\text{CV}}/\text{RMSE}_n$, and

$$\text{RMSE} = \sqrt{\frac{1}{M}\sum_{i=1}^{M}\left(\widehat{V}_i - \text{MSE}\right)^2},$$

and normalized biases

$$\text{NB} = 100 \times \left(\frac{1}{M}\sum_{i=1}^{M}\widehat{V}_i - \text{MSE}\right)/\text{MSE}$$

for all populations with $n \in \{100, 200, 500\}$ and $\sigma^2 \in \{0.01, 0.16\}$ under the design of simple random sampling without replacement. Apparently, absolute value of normalized bias is generally a decreasing function of sample size, and both RMSE ratios and normalized biases are stable across standard deviation. Furthermore, the CV variance estimator generally has smaller absolute value of normalized bias than the other variance estimators, and $\widehat{V}_g$ has smaller value than $\widehat{V}_n$, especially when sample size is small. In Linear3 and two Post-stratification cases, the RMSE ratios increase towards 1 as sample size increases, also the

CV variance estimator obviously has smaller RMSE than the other two estimators, and $\widehat{V}_g$ has smaller RMSE than $\widehat{V}_n$, especially when sample size is small. However, for the other population functions, there is no apparent difference of RMSE among the three variance estimators for all sample sizes and standard deviations.

Population	σ	n	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
1.Linear1	0.1	100	1.039	1.075	-1.296	0.338	3.911
		200	1.013	1.020	-1.868	-1.150	0.695
		500	1.001	0.990	-1.907	-1.710	-0.839
	0.4	100	1.039	1.075	-1.296	0.338	3.911
		200	1.013	1.020	-1.868	-1.150	0.695
		500	1.001	0.990	-1.907	-1.710	-0.839
2.Linear2	0.1	100	1.048	1.059	-7.144	-2.093	3.106
		200	1.005	1.002	-4.473	-2.095	0.691
		500	1.001	1.020	-1.257	-0.642	0.976
	0.4	100	1.048	1.059	-7.144	-2.093	3.106
		200	1.005	1.002	-4.473	-2.095	0.691
		500	1.001	1.020	-1.257	-0.642	0.976
3.Linear3	0.1	100	0.809	0.662	-22.587	-11.713	-2.661
		200	0.855	0.739	-11.992	-6.303	-1.445
		500	0.912	0.769	-7.118	-5.519	-2.651
	0.4	100	0.809	0.662	-22.587	-11.713	-2.661
		200	0.855	0.739	-11.992	-6.303	-1.445
		500	0.912	0.769	-7.118	-5.519	-2.651
4.Quadratic	0.1	100	1.020	1.017	-4.785	-2.296	0.894
		200	0.996	0.987	-3.196	-2.131	-0.368
		500	0.999	1.002	-1.019	-0.817	0.164
	0.4	100	1.034	1.056	-2.317	-0.661	2.527
		200	1.009	1.012	-2.225	-1.534	0.157
		500	1.000	1.001	-1.106	-0.957	-0.126
5.Collinear1	0.1	100	1.038	1.024	-7.197	-3.282	2.951
		200	1.014	0.988	-4.042	-2.209	1.059
		500	1.016	1.065	-0.254	0.215	1.920
	0.4	100	1.038	1.024	-7.197	-3.282	2.951
		200	1.014	0.988	-4.042	-2.209	1.059
		500	1.016	1.065	-0.254	0.215	1.920
6.Collinear2	0.1	100	1.039	1.024	-7.217	-3.265	2.951
		200	1.015	0.988	-4.034	-2.193	1.077
		500	1.017	1.066	-0.242	0.230	1.939
	0.4	100	1.039	1.024	-7.217	-3.265	2.951
		200	1.015	0.988	-4.034	-2.193	1.077
		500	1.017	1.066	-0.242	0.230	1.939
7.Collinear3	0.1	100	1.040	1.025	-7.226	-3.238	2.981

Continued...

Population	σ	n	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
		200	1.017	0.990	-4.017	-2.169	1.113
		500	1.018	1.067	-0.234	0.240	1.957
	0.4	100	1.040	1.025	-7.226	-3.238	2.981
		200	1.017	0.990	-4.017	-2.169	1.113
		500	1.018	1.067	-0.234	0.240	1.957
8.Collinear4	0.1	100	1.042	1.027	-7.228	-3.193	3.041
		200	1.019	0.992	-3.990	-2.130	1.170
		500	1.018	1.067	-0.233	0.244	1.970
	0.4	100	1.042	1.027	-7.228	-3.193	3.041
		200	1.019	0.992	-3.990	-2.130	1.170
		500	1.018	1.067	-0.233	0.244	1.970
9.Collinear5	0.1	100	1.046	1.031	-7.218	-3.061	3.194
		200	1.023	0.996	-3.945	-2.037	1.280
		500	1.019	1.065	-0.263	0.233	1.965
	0.4	100	1.046	1.031	-7.218	-3.061	3.194
		200	1.023	0.996	-3.945	-2.037	1.280
		500	1.019	1.065	-0.263	0.233	1.965
10.Post-ST1	0.1	100	0.982	0.845	-18.162	-10.307	1.222
		200	0.969	0.840	-10.786	-6.419	-0.857
		500	0.990	0.943	-4.035	-2.879	0.314
	0.4	100	0.982	0.845	-18.162	-10.307	1.222
		200	0.969	0.840	-10.786	-6.419	-0.857
		500	0.990	0.943	-4.035	-2.879	0.314
11.Post-ST2	0.1	100	0.908	0.676	-25.655	-17.701	-2.765
		200	0.909	0.733	-14.596	-9.345	-2.496
		500	0.958	0.836	-6.461	-5.024	-1.126
	0.4	100	0.908	0.676	-25.655	-17.701	-2.765
		200	0.909	0.733	-14.596	-9.345	-2.496
		500	0.958	0.836	-6.461	-5.024	-1.126

Table 2.3: Ratios of RMSEs of “g-corrected” and CV variance estimators to RMSE of “naive” variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from populations of size $N = 1000$ with $r = 0.5$.

The same overall behavior can be seen in the results of Table 2.3, which displays the ratios of RMSEs and normalized biases of three variance estimators for all populations under the design of stratified simple random sampling without replacement with $r = 0.5$.

Population	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
1.Linear1	0	1.009	1.012	-2.264	-2.009	0.261
	0.25	1.018	1.049	-0.010	0.461	2.600
	0.5	1.013	1.020	-1.868	-1.150	0.695
	0.75	1.022	1.025	-1.684	-0.463	0.980
	1	1.088	1.038	1.160	4.586	3.932
2.Linear2	0	1.015	1.020	-3.405	-2.464	1.561
	0.25	1.011	1.009	-3.770	-2.272	1.376
	0.5	1.005	1.002	-4.473	-2.095	0.691
	0.75	1.109	1.050	-2.870	1.664	2.412
	1	1.222	0.982	-7.206	3.477	-1.469
3.Linear3	0	0.957	0.889	-8.792	-6.547	2.594
	0.25	0.910	0.789	-10.376	-6.796	0.416
	0.5	0.855	0.739	-11.992	-6.303	-1.445
	0.75	0.766	0.675	-17.110	-7.871	-7.026
	1	0.742	0.768	-26.296	-5.756	-17.053
4.Quadratic	0	0.990	0.958	-4.267	-3.190	-1.372
	0.25	1.009	1.021	-1.917	-0.842	0.970
	0.5	0.996	0.987	-3.196	-2.131	-0.368
	0.75	1.006	1.021	-2.018	-0.910	0.871
	1	1.001	1.000	-2.492	-1.304	0.391
5.Collinear1	0	1.024	1.030	-3.610	-2.827	1.787
	0.25	1.053	1.073	-2.583	-1.450	2.799
	0.5	1.014	0.988	-4.042	-2.209	1.059
	0.75	1.055	1.069	-1.981	1.358	3.238
	1	1.242	0.995	-5.493	4.981	-0.807
6.Collinear2	0	1.024	1.029	-3.616	-2.815	1.776
	0.25	1.052	1.073	-2.583	-1.444	2.811
	0.5	1.015	0.988	-4.034	-2.193	1.077
	0.75	1.055	1.068	-2.039	1.319	3.176
	1	1.244	0.994	-5.514	4.967	-0.839
7.Collinear3	0	1.025	1.028	-3.635	-2.812	1.762
	0.25	1.052	1.074	-2.593	-1.449	2.831
	0.5	1.017	0.990	-4.017	-2.169	1.113
	0.75	1.055	1.067	-2.080	1.298	3.141
	1	1.245	0.993	-5.515	4.952	-0.851
8.Collinear4	0	1.025	1.027	-3.669	-2.816	1.742
	0.25	1.051	1.075	-2.615	-1.466	2.858
	0.5	1.019	0.992	-3.990	-2.130	1.170
	0.75	1.056	1.067	-2.098	1.304	3.140
	1	1.246	0.992	-5.496	4.937	-0.843
9.Collinear5	0	1.028	1.024	-3.765	-2.832	1.682

Continued...

Population	r	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
	0.25	1.050	1.077	-2.682	-1.507	2.910
	0.5	1.023	0.996	-3.945	-2.037	1.280
	0.75	1.058	1.069	-2.081	1.384	3.203
	1	1.243	0.991	-5.466	4.875	-0.825
10.Post-ST1	0	1.007	0.932	-8.901	-7.363	1.204
	0.25	1.006	0.950	-8.196	-5.631	1.957
	0.5	0.969	0.840	-10.786	-6.419	-0.857
	0.75	0.973	0.846	-12.265	-5.079	-2.178
	1	0.920	0.842	-21.805	-7.182	-12.255
11.Post-ST2	0	0.999	0.912	-10.340	-8.621	2.299
	0.25	0.984	0.891	-10.391	-7.213	2.204
	0.5	0.909	0.733	-14.596	-9.345	-2.496
	0.75	0.870	0.724	-18.019	-9.737	-6.016
	1	0.802	0.785	-29.153	-13.042	-18.127

Table 2.4: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from populations of size $N = 1000$ with model variance $\sigma^2 = 0.01$ and stratum sample sizes $(30, 40, 60, 70)$.

Table 2.4 displays simulation results under the design of stratified simple random sampling without replacement for all population functions with model variance $\sigma^2 = 0.01$ and the stratum sample sizes $n = (30, 40, 60, 70)$. Smaller value of r means that the relationship between the inclusion probability and the model error becomes weaker. In Linear3 and two Post-stratification cases, RMSE ratio of CV variance estimator decreases as r increases, and absolute value of normalized bias increases as r increases. However, for the other population functions, there is no obvious pattern.

2.5 Conclusion

In this chapter, we proposed a design-based CV variance estimator and compared it with other two variance estimators based on ratios of RMSEs and normalized biases with respect to MSE_p. We developed theoretical results by proving that the design-based properties of the CV variance estimator hold under some appropriate assumptions. By a simulation study, we showed that the CV variance estimator usually estimates the MSE_p quite well and works better than the other two variance estimators $\widehat{V}_n$ and $\widehat{V}_g$ in the cases of linear population with high dimensionality and post-stratification population with high number of strata, while being approximately as efficient for linear function with low dimension or multicollinearity, even with the misspecified case of quadratic population function. Hence, this design-based CV variance estimator can be used to estimate MSE_p in regression estimation.

CHAPTER 3. Shrinking for Linear Regression Estimation

3.1 Definition of the Estimator

The traditional variance estimator based on the usual linearization can not be used to choose between the π estimator and the regression estimator, because it will always pick the regression estimator. This will motivate the need to look at a different way to choose between the two estimators. In this chapter, a new estimator $\bar{y}_\gamma$ as linear combination of $\bar{y}_\pi$ and $\bar{y}_{\text{reg}}$ will be proposed:

$$\begin{aligned} \bar{y}_\gamma &= (1-\gamma)\,\bar{y}_\pi + \gamma \bar{y}_{\text{reg}} \\ &= \bar{y}_\pi + \gamma \left(\bar{\boldsymbol{X}}_N - \bar{\boldsymbol{X}}_\pi\right)^T \widehat{\boldsymbol{\beta}} \\ &= \frac{1}{N} \sum_{i \in s} g_{\gamma is} \frac{y_i}{\pi_i}, \end{aligned} \tag{3.1}$$

where $g_{\gamma is} = 1 + \gamma \left(\boldsymbol{t}_x - \widehat{\boldsymbol{t}}_{x\pi}\right)^T \widehat{\boldsymbol{T}}^{-1} \boldsymbol{x}_i / \sigma_i^2$, $\gamma \in [0,1]$ is a real valued fixed parameter. When $\gamma = 0$, $\bar{y}_\gamma$ will be the same as $\bar{y}_\pi$. While $\gamma = 1$, $\bar{y}_\gamma$ will be equal to $\bar{y}_{\text{reg}}$. If there is a strong linear relationship between $\boldsymbol{y}$ and $\boldsymbol{x}$, the regression estimator will have smaller error than the π estimator. Thus choosing γ close or equal to 1 will make the new estimator work better. On the contrary, when the π estimator works well, γ close to 0 will be a good choice for the proposed estimator. If the fitting of the model is very variable ($\boldsymbol{\beta}$ poorly estimated), then the added noise due to model fitting can "destroy" the improvement of efficiency due to regression estimation. Hence, the tuning constant γ can in principle serve to minimize the overall variance by trading off these two sources of variability. Thus, how to choose an appropriate γ value is important for the estimation. This will be discussed in next section.

By Taylor linearization, the proposed estimator $\bar{y}_\gamma$, is approximated by

$$\begin{aligned} \bar{y}_{\gamma,0} &= \bar{y}_\pi + \gamma\left(\bar{\boldsymbol{X}}_N - \bar{\boldsymbol{X}}_\pi\right)^T \boldsymbol{\beta}_N \\ &= \frac{1}{N}\gamma\sum_U y_i^0 + \frac{1}{N}\sum_s \check{E}_{\gamma i}, \end{aligned} \tag{3.2}$$

where $y_i^0 = \boldsymbol{x}_i^T\boldsymbol{\beta}_N$ and $\check{E}_{\gamma i} = E_{\gamma i}/\pi_i$ with $E_{\gamma i} = y_i - \gamma y_i^0$. The estimator $\bar{y}_\gamma$ is approximately unbiased for $\bar{y}_N$ with the approximate variance

$$AV\left(\bar{y}_\gamma\right) = \frac{1}{N^2}\sum_U\sum_U \Delta_{ij}\check{E}_{\gamma i}\check{E}_{\gamma j} \tag{3.3}$$

and the "naive" variance estimator

$$\widehat{V}\left(\bar{y}_\gamma\right) = \frac{1}{N^2}\sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\check{e}_{\gamma is}\check{e}_{\gamma js}, \tag{3.4}$$

where $\check{\Delta}_{ij} = \Delta_{ij}/\pi_{ij} = \left(\pi_{ij} - \pi_i\pi_j\right)/\pi_{ij}$, $\check{e}_{is} = e_{\gamma is}/\pi_i$ and $e_{\gamma is} = y_i - \gamma\boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}$.

Since the existence of γ in $g_{\gamma is}$, the estimator $\bar{y}_\gamma$ can not be expressed as a form similar to (2.12). Therefore, we fail to obtain a "g-corrected" variance estimator similar to (2.16).

If we use the linearized expression (3.2) for the estimator, all the extra variance due to estimating $\boldsymbol{\beta}_N$ disappears, which is why $\gamma = 1$ always wins if we just use expressions (3.3) or (3.4).

Here, we introduce a cross-validation variance estimator:

$$\widehat{V}_{\mathrm{CV}}\left(\bar{y}_\gamma\right) = \frac{1}{N^2}\sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\check{e}_{\gamma is}^{(-)}\check{e}_{\gamma js}^{(-)}, \tag{3.5}$$

where $\check{e}_{\gamma is}^{(-)} = e_{\gamma is}^{(-)}/\pi_i$, $e_{\gamma is}^{(-)} = y_i - \gamma\widehat{y}_i^{(-)}$ and $\widehat{y}_i = \boldsymbol{x}_i^T\widehat{\boldsymbol{\beta}}$. In this expression, we replace $\widehat{y}_i$ by the "leave-one-out" estimator $\widehat{y}_i^{(-)}$. Rewrite $\widehat{\boldsymbol{T}}$ and $\widehat{\boldsymbol{t}}$ as follows:

$$\widehat{\boldsymbol{T}} = \boldsymbol{X}_s^T\boldsymbol{V}_s^{-1}\boldsymbol{X}_s, \quad \text{and} \quad \widehat{\boldsymbol{t}} = \boldsymbol{X}_s^T\boldsymbol{V}_s^{-1}\boldsymbol{Y}_s,$$

where $\boldsymbol{X}_s$ represents the sample matrix of dimension $n_N \times J$, and $\boldsymbol{V}_s = \mathrm{diag}_{i\in s}\left\{\sigma_i^2\pi_i\right\}$ is $n_N \times n_N$ diagonal matrix. $\boldsymbol{Y}_s$ denotes the column vector of response values y_i for $i \in s$. Then by (2.17), we can show that

$$\begin{aligned} e_{\gamma is}^{(-)} &= y_i - \gamma\widehat{y}_i^{(-)} \\ &= \frac{y_i - \gamma\widehat{y}_i - (1-\gamma)\,w_{sii}^* y_i}{1 - w_{sii}^*} \\ &= \frac{e_{\gamma is}}{1 - w_{sii}^*} - \frac{(1-\gamma)\,w_{sii}^* y_i}{1 - w_{sii}^*}. \end{aligned} \tag{3.6}$$

Thereby $\widehat{V}_{\mathrm{CV}}\left(\bar{y}_\gamma\right)$ can be expressed as:

$$\widehat{V}_{\mathrm{CV}}\left(\bar{y}_\gamma\right)=\frac{1}{N^2}\sum_{i\in s}\sum_{j\in s}\frac{\check{\Delta}_{ij}}{\pi_i\pi_j}\left[\frac{e_{\gamma is}}{1-w^*_{sii}}-\frac{\left(1-\gamma\right)w^*_{sii}y_i}{1-w^*_{sii}}\right]\left[\frac{e_{\gamma js}}{1-w^*_{sjj}}-\frac{\left(1-\gamma\right)w^*_{sjj}y_j}{1-w^*_{sjj}}\right]. \quad (3.7)$$

3.2 Theoretical Properties

For simplicity, we will only consider the case with the sample size, denoted by n_N, fixed for each N, and also assume that $n_N \to \infty$. As above, J is fixed. In order to prove our theoretical results, we make the following technical assumptions.

A1. (Sampling rate $n_N N^{-1}$). As $N\to\infty$, $n_N N^{-1}\to\pi\in(0,1)$.

A2. (Inclusion probabilities π_i and π_j). For all N, $\min_{i\in U}\pi_i\geq\lambda>0$, $\min_{i,j\in U}\pi_{ij}\geq\lambda^*>0$, $\lim_{N\to\infty} n_N \max_{i,j\in U, i\neq j}\left|\Delta_{ij}\right|<\infty$.

A3. Assume that $\left(\frac{n_N}{N^2}\widehat{\boldsymbol{T}}\right)^{-1}$ and $\left(\frac{n_N}{N^2}\boldsymbol{T}\right)^{-1}$ exist for all samples.

A4. $\max_{i\in U}\left|y_i\right|<C_y$, and $\max_{i\in U, j\in\{1,\ldots,J\}}\left|x_{ij}\right|<C_x$, where C_y and C_x are some positive constants.

A5. $0<\sigma_L\leq\min_{i\in U}\sigma_i\leq\max_{i\in U}\sigma_i\leq\sigma_U<\infty$, where σ_L and σ_U are some positive constants.

A6. Additional assumptions involving higher-order inclusion probabilities:

$$\lim_{N\to\infty}\max_{(i,j,k)\in D_{3,N}}\left|\mathrm{E}_p\left[\left(I_i-\pi_i\right)\left(I_jI_k-\pi_{jk}\right)\right]\right|=0$$

$$\lim_{N\to\infty}\max_{(i,j,k,l)\in D_{4,N}}\left|\mathrm{E}_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_kI_l-\pi_{kl}\right)\right]\right|=0$$

$$\lim_{N\to\infty} n_N\max_{(i,j,k)\in D_{3,N}}\left|\mathrm{E}_p\left[\left(I_i-\pi_i\right)^2\left(I_j-\pi_j\right)\left(I_k-\pi_k\right)\right]\right|<\infty$$

$$\lim_{N\to\infty} n_N^2\max_{(i,j,k,l)\in D_{4,N}}\left|\mathrm{E}_p\left[\left(I_i-\pi_i\right)\left(I_j-\pi_j\right)\left(I_k-\pi_k\right)\left(I_l-\pi_l\right)\right]\right|<\infty$$

where $D_{t,N}$ denotes the set of all distinct t-tuples from U.

Assumption A3 ensures that $\widehat{\boldsymbol{\beta}}$ and $\boldsymbol{\beta}_N$ exist. The following results establish design consistency of variance estimator of $\bar{y}_\gamma$.

Theorem 3.2.1. *Let assumptions A1-A6 hold. Then we have the following result:*

$$\sup_{\gamma\in[0,1]}\left|\mathrm{MSE}_p\left(\bar{y}_\gamma\right)-\frac{1}{N^2}\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma\boldsymbol{x}_i^T\boldsymbol{\beta}_N}{\pi_i}\frac{y_j-\gamma\boldsymbol{x}_j^T\boldsymbol{\beta}_N}{\pi_j}\right| \quad = \quad o\left(n_N^{-1}\right).$$

Proof of Theorem 3.2.1: Since

$$\begin{aligned}\bar{y}_\gamma-\bar{y}_N &= \frac{1}{N}\sum_U y_i\left(\frac{I_i}{\pi_i}-1\right)+\frac{\gamma}{N}\sum_U\widehat{y}_i\left(1-\frac{I_i}{\pi_i}\right)\\ &= \frac{1}{N}\sum_U\left(y_i-\gamma\widehat{y}_i\right)\left(\frac{I_i}{\pi_i}-1\right).\end{aligned}$$

Let

$$a_{\gamma,N}=\sum_U\left(y_i-\gamma y_i^0\right)\left(\frac{I_i}{\pi_i}-1\right)\quad\text{and}\quad b_{\gamma,N}=\gamma\sum_U\left(y_i^0-\widehat{y}_i\right)\left(\frac{I_i}{\pi_i}-1\right).$$

Then, by assumptions A1, A2 and Lemma A.2.4,

$$\begin{aligned}\mathrm{E}_p a_{\gamma,N}^2 &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\\ &\le \frac{1}{\lambda^2}\sum_{i\in U}\max_{i\in U}\left|\pi_i\left(1-\pi_i\right)\right|\left\{\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\right\}^2\\ &\quad+\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\left\{\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\right\}^2\\ &= O\left(N\right)+O\left(\frac{N^2}{n_N}\right)\\ &= O\left(\frac{N^2}{n_N}\right),\end{aligned}$$

Then it follows that

$$\sup_{\gamma\in[0,1]}\frac{1}{N^2}\mathrm{E}_p\left[a_{\gamma,N}^2\right]=O\left(\frac{1}{n_N}\right).$$

As the fact that $|\gamma|\le 1$, by assumptions A1, A2, A4, A6 and Lemma A.2.2, it can be shown that

$$\sup_{\gamma\in[0,1]}\frac{1}{N^2}\mathrm{E}_p\left[b_{\gamma,N}^2\right]=o\left(\frac{1}{n_N}\right).$$

The proof is similar to the proof of Theorem 2.3.1. Then, we will have that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} \frac{1}{N^2}\mathrm{E}_p\left[\left|a_{\gamma,N}b_{\gamma,N}\right|\right] &\leq \sqrt{\sup_{\gamma\in[0,1]} \frac{1}{N^2}\mathrm{E}_p\left[a_{\gamma,N}^2\right] \sup_{\gamma\in[0,1]} \frac{1}{N^2}\mathrm{E}_p\left[b_{\gamma,N}^2\right]} \\
&= \sqrt{O\left(\frac{1}{n_N}\right) o\left(\frac{1}{n_N}\right)} \\
&= o\left(\frac{1}{n_N}\right).
\end{aligned}$$

Thereby,

$$\begin{aligned}
\sup_{\gamma\in[0,1]} \left|\mathrm{MSE}_p\left(\bar{y}_\gamma\right) - \frac{1}{N^2}\mathrm{E}_p\left[a_{\gamma,N}^2\right]\right| &= \sup_{\gamma\in[0,1]} \frac{1}{N^2}\left\{\left|\mathrm{E}_p\left[b_{\gamma,N}^2\right] + 2\mathrm{E}_p\left[a_{\gamma,N}b_{\gamma,N}\right]\right|\right\} \\
&\leq \sup_{\gamma\in[0,1]} \frac{1}{N^2}\mathrm{E}_p\left[b_{\gamma,N}^2\right] + \sup_{\gamma\in[0,1]} \frac{2}{N^2}\mathrm{E}_p\left[\left|a_{\gamma,N}b_{\gamma,N}\right|\right] \\
&= o\left(\frac{1}{n_N}\right) + o\left(\frac{1}{n_N}\right) \\
&= o\left(\frac{1}{n_N}\right).
\end{aligned}$$

Therefore the result follows.

□

Theorem 3.2.2. *Under assumptions A1-A6, we have that the estimator*

$$\sup_{\gamma\in[0,1]} \left|\widehat{V}_{\mathrm{CV}}\left(\bar{y}_\gamma\right) - \frac{1}{N^2}\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i - \gamma\boldsymbol{x}_i^T\boldsymbol{\beta}_N}{\pi_i}\frac{y_j - \gamma\boldsymbol{x}_j^T\boldsymbol{\beta}_N}{\pi_j}\right| = o_p\left(n_N^{-1}\right).$$

Proof of Theorem 3.2.2: First we separate the expression into three terms as follows

$$\begin{aligned}
\widehat{V}_{\mathrm{CV}}\left(\bar{y}_\gamma\right) &= \frac{1}{N^2}\sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\frac{y_i - \gamma\widehat{y}_i}{\pi_i}\frac{y_j - \gamma\widehat{y}_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*} \\
&\quad -\frac{2}{N^2}\sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\frac{y_i - \gamma\widehat{y}_i}{\pi_i}\frac{(1-\gamma)\,y_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{w_{sjj}^*}{1-w_{sjj}^*} \\
&\quad +\frac{1}{N^2}\sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\frac{(1-\gamma)\,y_i}{\pi_i}\frac{(1-\gamma)\,y_j}{\pi_j}\frac{w_{sii}^*}{1-w_{sii}^*}\frac{w_{sjj}^*}{1-w_{sjj}^*} \\
&= \frac{1}{N^2}\left(V_a - 2V_b + V_c\right).
\end{aligned}$$

Then, for V_a, we could rewrite it as

$$\begin{aligned} V_a &= \sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*} \\ &\quad +\sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\frac{\gamma\left(y_i^0-\widehat{y}_i\right)}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*} \\ &\quad +2\sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*} \\ &= V_{a1}+V_{a2}+2V_{a3}. \end{aligned}$$

Furthermore, V_{a1} can be written as

$$\begin{aligned} V_{a1} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*} \\ &\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}\left(\frac{I_{ij}}{\pi_{ij}}-1\right) \\ &= V_{a11}+V_{a12}, \end{aligned}$$

where

$$\begin{aligned} V_{a11} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j} \\ &\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}-1\right) \\ &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}+V_{a11}^*. \end{aligned}$$

Suppose assumptions A1 and A2 hold, from Lemma A.2.4 and A.2.6, it follows that

$$\begin{aligned}
\sup_{\gamma\in[0,1]}|V_{a11}^{*}| &\leq \sup_{\gamma\in[0,1]}\sum_{i\in U}|\pi_i(1-\pi_i)|\frac{\left(y_i-\gamma y_i^0\right)^2}{\pi_i^2}\left|g_{ii}^{*}\left(\widehat{\boldsymbol{T}}\right)\right| \\
&\quad+\sup_{\gamma\in[0,1]}\sum_{(i,j)\in D_{2,N}}|\Delta_{ij}|\left|\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\right|\left|g_{ij}^{*}\left(\widehat{\boldsymbol{T}}\right)\right| \\
&\leq \frac{1}{\lambda^2}\sum_{i\in U}\left\{\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\right\}^2\max_{i,j\in U}\left|g_{ij}^{*}\left(\widehat{\boldsymbol{T}}\right)\right| \\
&\quad+\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}|\Delta_{ij}|\left\{\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\right\}^2\max_{i,j\in U}\left|g_{ij}^{*}\left(\widehat{\boldsymbol{T}}\right)\right| \\
&= O(1)+O\left(\frac{N}{n_N}\right) \\
&= O\left(\frac{N}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U}|\Delta_{ii}|=\max_{i\in U}|\pi_i(1-\pi_i)|\leq 1$. Rewrite V_{a12} as follows,

$$\begin{aligned}
V_{a12} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\left(\frac{I_{ij}}{\pi_{ij}}-1\right) \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\left(\frac{1}{1-w_{sii}^{*}}\frac{1}{1-w_{sjj}^{*}}-1\right)\left(\frac{I_{ij}}{\pi_{ij}}-1\right) \\
&= V_{a121}+V_{a122}.
\end{aligned}$$

Where, $\mathrm{E}_p[V_{a121}]=0$, and

$$\begin{aligned}
\mathrm{Var}_p[V_{a121}] &= E_p\left[V_{a121}^2\right] \\
&= \sum_{i,j,k,l\in U}\Delta_{ij}\Delta_{kl}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\frac{y_k-\gamma y_k^0}{\pi_k}\frac{y_l-\gamma y_l^0}{\pi_l}\mathrm{E}_p\left[\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\frac{I_kI_l}{\pi_{kl}}-1\right)\right].
\end{aligned}$$

Then, by assumptions A1, A2, A6 and Lemma A.2.4,

$$\begin{aligned}
\sup_{\gamma\in[0,1]} \left|\mathrm{Var}_p\left[V_{a121}\right]\right| &\leq \frac{1}{\lambda^4}\sum_{i,j,k,l\in U}\max_{i,j\in U}\left|\Delta_{ij}\right|\max_{k,l\in U}\left|\Delta_{kl}\right|\left\{\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\right\}^4 \\
&\quad\times\max_{i,j,k,l\in U}\left|E_p\left[\left(\frac{I_iI_j}{\pi_{ij}}-1\right)\left(\frac{I_kI_l}{\pi_{kl}}-1\right)\right]\right| \\
&= \frac{2}{\lambda^5\lambda^*}\sum_{(i,j,k)\in D_{3,N}} O\left(\frac{1}{n_N}\right)\max_{(i,j,k)\in D_{3,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_k-\pi_k\right)\right]\right| \\
&\quad+\frac{4}{\lambda^4\lambda^{*2}}\sum_{(i,j,k)\in D_{3,N}} O\left(\frac{1}{n_N^2}\right)\max_{(i,j,k)\in D_{3,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_iI_k-\pi_{ik}\right)\right]\right| \\
&\quad+\frac{4}{\lambda^5\lambda^*}\sum_{(i,j)\in D_{2,N}} O\left(\frac{1}{n_N}\right)\max_{(i,j)\in D_{2,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_i-\pi_i\right)\right]\right| \\
&\quad+\frac{1}{\lambda^6}\sum_{i\in U} O\left(1\right)\max_{i\in U}\left|E_p\left[\left(I_i-\pi_i\right)^2\right]\right| \\
&\quad+\frac{1}{\lambda^4\lambda^{*2}}\sum_{(i,j,k,l)\in D_{4,N}} O\left(\frac{1}{n_N^2}\right)\max_{(i,j,k,l)\in D_{4,N}}\left|E_p\left[\left(I_iI_j-\pi_{ij}\right)\left(I_kI_l-\pi_{kl}\right)\right]\right| \\
&= o\left(\frac{N^3}{n_N}\right)+O\left(\frac{N^3}{n_N^2}\right)+O\left(\frac{N^2}{n_N}\right)+O\left(N\right)+o\left(\frac{N^4}{n_N^2}\right) \\
&= o\left(\frac{N^4}{n_N^2}\right),
\end{aligned}$$

which implies that $\sup_{\gamma\in[0,1]}\left|V_{a121}\right| = o_p\left(\frac{N^2}{n_N}\right)$.

By assumption A2,

$$\begin{aligned}
\max_{i\in U}\left|\frac{I_i}{\pi_i}-1\right| &= \max_{i\in U}\left(1,\frac{1}{\pi_i}-1\right) \\
&\leq \frac{1}{\lambda},
\end{aligned} \tag{3.8}$$

and,

$$\begin{aligned}
\max_{(i,j)\in D_{2,N}}\left|\frac{I_iI_j}{\pi_{ij}}-1\right| &= \max_{i\in U}\left(1,\frac{1}{\pi_{ij}}-1\right) \\
&\leq \frac{1}{\lambda^*}.
\end{aligned} \tag{3.9}$$

Then V_{a122} should have the same order as V^*_{a11}. Therefore, under assumptions A1 and A2, by Lemma A.2.4 and A.2.6, it can be shown that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{a122}| &\leq \sum_{i\in U} \frac{1}{\lambda^3} \left\{ \sup_{\gamma\in[0,1]} \max_{i\in U} \left|y_i - \gamma y_i^0\right| \right\}^2 \max_{i,j\in U} \left| g^*_{ij}\left(\widehat{\boldsymbol{T}}\right)\right| \\
&\quad + \sum_{(i,j)\in D_{2,N}} \frac{1}{\lambda^2\lambda^*} \max_{(i,j)\in D_{2,N}} |\Delta_{ij}| \left\{ \sup_{\gamma\in[0,1]} \max_{i\in U} \left|y_i - \gamma y_i^0\right| \right\}^2 \max_{i,j\in U} \left| g^*_{ij}\left(\widehat{\boldsymbol{T}}\right)\right| \\
&= O(1) + O\left(\frac{N}{n_N}\right) \\
&= O\left(\frac{N}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U} |\Delta_{ii}| = \max_{i\in U} |\pi_i (1-\pi_i)| \leq 1$. Thus, it follows that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{a12}| &\leq \sup_{\gamma\in[0,1]} |V_{a121}| + \sup_{\gamma\in[0,1]} |V_{a122}| \\
&= o_p\left(\frac{N^2}{n_N}\right) + O\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}$$

Next, we will show that

$$\sup_{\gamma\in[0,1]} |V_{a2}| = o_p\left(\frac{N^2}{n_N}\right).$$

Note that

$$\begin{aligned}
V_{a2} &= \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{\gamma\left(y_i^0 - \widehat{y}_i\right)}{\pi_i} \frac{\gamma\left(y_j^0 - \widehat{y}_j\right)}{\pi_j} \frac{1}{1-w^*_{sii}} \frac{1}{1-w^*_{sjj}} \\
&\quad + \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{\gamma\left(y_i^0 - \widehat{y}_i\right)}{\pi_i} \frac{\gamma\left(y_j^0 - \widehat{y}_j\right)}{\pi_j} \frac{1}{1-w^*_{sii}} \frac{1}{1-w^*_{sjj}} \left(\frac{I_{ij}}{\pi_{ij}} - 1\right) \\
&= V_{a21} + V_{a22}.
\end{aligned}$$

And,

$$\begin{aligned}
V_{a21} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\gamma\left(y_i^0-\widehat{y}_i\right)}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j} \\
&\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\gamma\left(y_i^0-\widehat{y}_i\right)}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}-1\right) \\
&= \gamma^2\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right) \\
&\quad +\gamma^2\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right) \\
&= V_{a211}+V_{a212}.
\end{aligned}$$

Suppose assumptions A1, A2 and A4 hold, by Lemma A.2.1, it can be shown that

$$\begin{aligned}
\sup_{\gamma\in[0,1]}\left|V_{a211}\right| &\le \frac{1}{\lambda^2}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right|^T\sum_{i\in U}\left|\pi_i\left(1-\pi_i\right)\boldsymbol{x}_i\boldsymbol{x}_i^T\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&\quad +\frac{1}{\lambda^2}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right|^T\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\left|\boldsymbol{x}_i\boldsymbol{x}_j^T\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&= o_p(1)\,O(N)\,o_p(1)+o_p(1)\,O\left(\frac{N^2}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N^2}{n_N}\right),
\end{aligned}$$

and by Lemma A.2.6

$$\begin{aligned}
\sup_{\gamma\in[0,1]}\left|V_{a212}\right| &\le \frac{1}{\lambda^2}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right|^T\sum_{i\in U}\left|\pi_i\left(1-\pi_i\right)\boldsymbol{x}_i\boldsymbol{x}_i^T\right|\max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&\quad +\frac{1}{\lambda^2}\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right|^T\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\left|\boldsymbol{x}_i\boldsymbol{x}_j^T\right| \\
&\quad \times\max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&= o_p(1)\,O(1)\,o_p(1)+o_p(1)\,O\left(\frac{N}{n_N}\right)o_p(1) \\
&= o_p\left(\frac{N}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U}|\Delta_{ii}| = \max_{i\in U}|\pi_i(1-\pi_i)| \le 1$. Then, it follows that

$$\begin{aligned}
\sup_{\gamma\in[0,1]}|V_{a21}| &\le \sup_{\gamma\in[0,1]}|V_{a211}| + \sup_{\gamma\in[0,1]}|V_{a212}| \\
&= o_p\left(\frac{N^2}{n_N}\right) + o_p\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}$$

Rewrite V_{a22} as follows,

$$\begin{aligned}
V_{a22} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\gamma\left(y_i^0-\widehat{y}_i\right)}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\left(\frac{I_{ij}}{\pi_{ij}}-1\right) \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\gamma\left(y_i^0-\widehat{y}_i\right)}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}-1\right)\left(\frac{I_{ij}}{\pi_{ij}}-1\right) \\
&= \gamma^2\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\frac{I_{ij}}{\pi_{ij}}-1\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right) \\
&\quad+\gamma^2\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right)^T\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{\boldsymbol{x}_i}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\frac{I_{ij}}{\pi_{ij}}-1\right)\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right) \\
&= V_{a221}+V_{a222}.
\end{aligned}$$

From (3.8) and (3.9), V_{a22} should have the same order as V_{a21}. It follows that

$$\sup_{\gamma\in[0,1]}|V_{a22}| = o_p\left(\frac{N^2}{n_N}\right).$$

Therefore, it can be shown that

$$\begin{aligned}
\sup_{\gamma\in[0,1]}|V_{a2}| &\le \sup_{\gamma\in[0,1]}|V_{a21}| + \sup_{\gamma\in[0,1]}|V_{a22}| \\
&= o_p\left(\frac{N^2}{n_N}\right) + o_p\left(\frac{N^2}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}$$

Similarly, we write V_{a3} as follows,

$$\begin{aligned}
V_{a3} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*} \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}\left(\frac{I_{ij}}{\pi_{ij}}-1\right) \\
&= V_{a31}+V_{a32}.
\end{aligned}$$

And note,

$$\begin{aligned}
V_{a31} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j} \\
&\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}-1\right) \\
&= \gamma\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right) \\
&\quad +\gamma\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right) \\
&= V_{a311}+V_{a312}.
\end{aligned}$$

Suppose assumptions A1, A2 and A4 hold, from Lemma A.2.1 and A.2.4 we can show that

$$\begin{aligned}
\sup_{\gamma\in[0,1]}\left|V_{a311}\right| &= \sup_{\gamma\in[0,1]}\left|\gamma\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right)\right| \\
&\leq \frac{1}{\lambda^2}\sum_{i\in U}\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\max_{i\in U}\left|\boldsymbol{x}_i\right|^T\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&\quad +\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\max_{j\in U}\left|\boldsymbol{x}_j\right|^T\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&= O\left(N\right)o_p\left(1\right)+O\left(\frac{N^2}{n_N}\right)o_p\left(1\right) \\
&= o_p\left(\frac{N^2}{n_N}\right),
\end{aligned}$$

and,

$$\begin{aligned}
\sup_{\gamma\in[0,1]}\left|V_{a312}\right| &= \sup_{\gamma\in[0,1]}\left|\gamma\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right)\right| \\
&\leq \frac{1}{\lambda^2}\sum_{i\in U}\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\max_{i\in U}\left|\boldsymbol{x}_i\right|^T\max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&\quad +\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}\left|\Delta_{ij}\right|\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\max_{j\in U}\left|\boldsymbol{x}_j\right|^T \\
&\quad \times\max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right|\left|\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_N\right| \\
&= O\left(1\right)o_p\left(1\right)+O\left(\frac{N}{n_N}\right)o_p\left(1\right) \\
&= o_p\left(\frac{N}{n_N}\right),
\end{aligned}$$

as the fact that $\max_{i\in U}|\Delta_{ii}| = \max_{i\in U}|\pi_i(1-\pi_i)| \le 1$. Thereby,

$$
\begin{aligned}
\sup_{\gamma\in[0,1]}|V_{a31}| &\le \sup_{\gamma\in[0,1]}|V_{a311}| + \sup_{\gamma\in[0,1]}|V_{a312}| \\
&= o_p\left(\frac{N^2}{n_N}\right) + o_p\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}
$$

Write V_{a32} as follows,

$$
\begin{aligned}
V_{a32} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\left(\frac{I_{ij}}{\pi_{ij}}-1\right) \\
&\quad+\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\gamma\left(y_j^0-\widehat{y}_j\right)}{\pi_j}\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}-1\right)\left(\frac{I_{ij}}{\pi_{ij}}-1\right) \\
&= \gamma\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}\left(\frac{I_{ij}}{\pi_{ij}}-1\right)\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right) \\
&\quad+\gamma\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{\boldsymbol{x}_j^T}{\pi_j}g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\left(\frac{I_{ij}}{\pi_{ij}}-1\right)\left(\boldsymbol{\beta}_N-\widehat{\boldsymbol{\beta}}\right) \\
&= V_{a321}+V_{a322}.
\end{aligned}
$$

Similarly, by (3.8) and (3.9), V_{a32} should have the same order as V_{a31}. It follows that

$$
\begin{aligned}
\sup_{\gamma\in[0,1]}|V_{a3}| &\le \sup_{\gamma\in[0,1]}|V_{a31}| + \sup_{\gamma\in[0,1]}|V_{a32}| \\
&= o_p\left(\frac{N^2}{n_N}\right) + o_p\left(\frac{N^2}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}
$$

Therefore,

$$
\begin{aligned}
\sup_{\gamma\in[0,1]}\left|V_a-\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{E_{\gamma i}}{\pi_i}\frac{E_{\gamma j}}{\pi_j}\right| &\le \sup_{\gamma\in[0,1]}|V_{a11}^*| + \sup_{\gamma\in[0,1]}|V_{a12}| + \sup_{\gamma\in[0,1]}|V_{a2}| + \sup_{\gamma\in[0,1]}|V_{a3}| \\
&= O\left(\frac{N}{n_N}\right) + o_p\left(\frac{N^2}{n_N}\right) + o_p\left(\frac{N^2}{n_N}\right) + 2o_p\left(\frac{N^2}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}
$$

Secondly, we will show that $\sup_{\gamma\in[0,1]}|V_b| = o_p\left(\frac{N^2}{n_N}\right)$. The proof is very similar to the above. First, we rewrite V_b as

$$\begin{aligned} V_b &= \sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{(1-\gamma)\,y_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{w_{sjj}^*}{1-w_{sjj}^*} \\ &\quad +\gamma\sum_{i\in s}\sum_{j\in s}\check{\Delta}_{ij}\frac{y_i^0-\widehat{y}_i}{\pi_i}\frac{(1-\gamma)\,y_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{w_{sjj}^*}{1-w_{sjj}^*} \\ &= V_{b1}+V_{b2}. \end{aligned}$$

And,

$$\begin{aligned} V_{b1} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{(1-\gamma)\,y_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{w_{sjj}^*}{1-w_{sjj}^*} \\ &\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{(1-\gamma)\,y_j}{\pi_j}\frac{1}{1-w_{sii}^*}\frac{w_{sjj}^*}{1-w_{sjj}^*}\left(\frac{I_{ij}}{\pi_{ij}}-1\right) \\ &= V_{b11}+V_{b12}. \end{aligned}$$

Then,

$$\begin{aligned} V_{b11} &= \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{(1-\gamma)\,y_j}{\pi_j}w_{sjj}^* \\ &\quad +\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{(1-\gamma)\,y_j}{\pi_j}w_{sjj}^*\left(\frac{1}{1-w_{sii}^*}\frac{1}{1-w_{sjj}^*}-1\right) \\ &= V_{b111}+V_{b112}. \end{aligned}$$

By assumptions A1, A2, and Lemma A.2.4, A.2.5 and A.2.6 we will have the following results,

$$\begin{aligned} \sup_{\gamma\in[0,1]}|V_{b111}| &\leq \sup_{\gamma\in[0,1]}\sum_{i\in U}|\pi_i(1-\pi_i)|\left|\frac{(1-\gamma)\,y_i\left(y_i-\gamma y_i^0\right)}{\pi_i^2}\right||w_{sii}^*| \\ &\quad +\sup_{\gamma\in[0,1]}\sum_{(i,j)\in D_{2,N}}|\Delta_{ij}|\left|\frac{y_i-\gamma y_i^0}{\pi_i}\frac{(1-\gamma)\,y_j}{\pi_j}\right||w_{sii}^*| \\ &\leq \frac{1}{\lambda^2}\sum_{i\in U}\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\max_{i\in U}|y_i|\max_{i\in U}|w_{sii}^*| \\ &\quad +\frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}}\max_{(i,j)\in D_{2,N}}|\Delta_{ij}|\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i-\gamma y_i^0\right|\max_{j\in U}|y_j|\max_{j\in U}\left|w_{sjj}^*\right| \\ &= O(1)+O\left(\frac{N}{n_N}\right) \\ &= O\left(\frac{N}{n_N}\right), \end{aligned}$$

and,

$$
\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{b112}| &\leq \sup_{\gamma\in[0,1]} \sum_{i\in U} |\pi_i (1-\pi_i)| \left| \frac{(1-\gamma)\, y_i \left(y_i - \gamma y_i^0\right)}{\pi_i^2} \right| |w_{sii}^*| \left| g_{ii}^* \left(\widehat{\boldsymbol{T}}\right) \right| \\
&\quad + \sup_{\gamma\in[0,1]} \sum_{(i,j)\in D_{2,N}} |\Delta_{ij}| \left| \frac{y_i - \gamma y_i^0}{\pi_i} \frac{(1-\gamma)\, y_j}{\pi_j} \right| |w_{sii}^*| \left| g_{ij}^* \left(\widehat{\boldsymbol{T}}\right) \right| \\
&\leq \frac{1}{\lambda^2} \sum_{i\in U} \sup_{\gamma\in[0,1]} \max_{i\in U} \left| y_i - \gamma y_i^0 \right| \max_{i\in U} |y_i| \max_{i\in U} |w_{sii}^*| \max_{i,j\in U} \left| g_{ij}^* \left(\widehat{\boldsymbol{T}}\right) \right| \\
&\quad + \frac{1}{\lambda^2} \sum_{(i,j)\in D_{2,N}} \max_{(i,j)\in D_{2,N}} |\Delta_{ij}| \sup_{\gamma\in[0,1]} \max_{i\in U} \left| y_i - \gamma y_i^0 \right| \max_{j\in U} |y_j| \max_{j\in U} \left| w_{sjj}^* \right| \\
&\quad \times \max_{i,j\in U} \left| g_{ij}^* \left(\widehat{\boldsymbol{T}}\right) \right| \\
&= O\left(\frac{1}{N}\right) + O\left(\frac{1}{n_N}\right) \\
&= O\left(\frac{1}{n_N}\right).
\end{aligned}
$$

Thus, we can show that

$$
\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{b11}| &\leq \sup_{\gamma\in[0,1]} |V_{b111}| + \sup_{\gamma\in[0,1]} |V_{b112}| \\
&= O\left(\frac{N}{n_N}\right) + O\left(\frac{1}{n_N}\right) \\
&= O\left(\frac{N}{n_N}\right).
\end{aligned}
$$

Similarly, by (3.8) and (3.9), V_{b12} has the same order as V_{b11}. Then, it follows that

$$
\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{b1}| &\leq \sup_{\gamma\in[0,1]} |V_{b11}| + \sup_{\gamma\in[0,1]} |V_{b12}| \\
&= O\left(\frac{N}{n_N}\right) + O\left(\frac{N}{n_N}\right) \\
&= O\left(\frac{N}{n_N}\right).
\end{aligned}
$$

Then, we will express V_{b2} as follows

$$
\begin{aligned}
V_{b2} &= \gamma \sum_{i\in U} \sum_{j\in U} \Delta_{ij} \frac{y_i^0 - \widehat{y}_i}{\pi_i} \frac{(1-\gamma)\, y_j}{\pi_j} \frac{1}{1 - w_{sii}^*} \frac{w_{sjj}^*}{1 - w_{sjj}^*} \\
&\quad + \gamma \sum_{i\in U} \sum_{j\in U} \Delta_{ij} \frac{y_i^0 - \widehat{y}_i}{\pi_i} \frac{(1-\gamma)\, y_j}{\pi_j} \frac{1}{1 - w_{sii}^*} \frac{w_{sjj}^*}{1 - w_{sjj}^*} \left(\frac{I_{ij}}{\pi_{ij}} - 1 \right) \\
&= V_{b21} + V_{b22},
\end{aligned}
$$

where,

$$\begin{aligned}
V_{b21} &= \gamma \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{y_i^0 - \widehat{y}_i}{\pi_i} \frac{(1-\gamma)\, y_j}{\pi_j} w_{sjj}^* \\
&\quad + \gamma \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{y_i^0 - \widehat{y}_i}{\pi_i} \frac{(1-\gamma)\, y_j}{\pi_j} w_{sjj}^* \left(\frac{1}{1-w_{sii}^*} \frac{1}{1-w_{sjj}^*} - 1 \right) \\
&= \gamma (1-\gamma) \left(\boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}} \right)^T \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{\boldsymbol{x}_i}{\pi_i} \frac{y_j}{\pi_j} w_{sjj}^* \\
&\quad + \gamma (1-\gamma) \left(\boldsymbol{\beta}_N - \widehat{\boldsymbol{\beta}} \right)^T \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{\boldsymbol{x}_i}{\pi_i} \frac{y_j}{\pi_j} w_{sjj}^* g_{ij}^* \left(\widehat{\boldsymbol{T}} \right) \\
&= V_{b211} + V_{b212}.
\end{aligned}$$

Suppose assumptions A1, A2 and A4 hold, by Lemma A.2.1 and A.2.5, it can be shown that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{b211}| &\leq \frac{1}{\lambda^2} \left| \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_N \right|^T \sum_{i\in U} \left| \pi_i (1-\pi_i) \boldsymbol{x}_i y_i \right| \max_{i\in U} |w_{sii}^*| \\
&\quad + \frac{1}{\lambda^2} \left| \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_N \right|^T \sum_{(i,j)\in D_{2,N}} \max_{(i,j)\in D_{2,N}} |\Delta_{ij}| \, |\boldsymbol{x}_i y_j| \max_{j\in U} \left| w_{sjj}^* \right| \\
&= o_p(1)\, O(1) + o_p(1)\, O\left(\frac{N}{n_N} \right) \\
&= o_p\left(\frac{N}{n_N} \right),
\end{aligned}$$

and by Lemma A.2.6,

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{b212}| &\leq \frac{1}{\lambda^2} \left| \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_N \right|^T \sum_{i\in U} \left| \pi_i (1-\pi_i) \boldsymbol{x}_i y_i \right| \max_{i\in U} |w_{sii}^*| \max_{i,j\in U} \left| g_{ij}^* \left(\widehat{\boldsymbol{T}} \right) \right| \\
&\quad + \frac{1}{\lambda^2} \left| \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_N \right|^T \sum_{(i,j)\in D_{2,N}} \max_{(i,j)\in D_{2,N}} |\Delta_{ij}| \, |\boldsymbol{x}_i y_j| \max_{j\in U} \left| w_{sjj}^* \right| \max_{i,j\in U} \left| g_{ij}^* \left(\widehat{\boldsymbol{T}} \right) \right| \\
&= o_p(1)\, O\left(\frac{1}{N} \right) + o_p(1)\, O\left(\frac{1}{n_N} \right) \\
&= o_p\left(\frac{1}{n_N} \right).
\end{aligned}$$

Then, it follows that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{b21}| &\leq \sup_{\gamma\in[0,1]} |V_{b211}| + \sup_{\gamma\in[0,1]} |V_{b212}| \\
&= o_p\left(\frac{N}{n_N} \right) + o_p\left(\frac{1}{n_N} \right) \\
&= o_p\left(\frac{N}{n_N} \right).
\end{aligned}$$

In the same way, by (3.8) and (3.9), V_{b22} has the same order as V_{b21}, which implies that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{b2}| &\leq \sup_{\gamma\in[0,1]} |V_{b21}| + \sup_{\gamma\in[0,1]} |V_{b22}| \\
&= o_p\left(\frac{N}{n_N}\right) + o_p\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N}{n_N}\right).
\end{aligned}$$

Then, it follows that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_b| &\leq \sup_{\gamma\in[0,1]} |V_{b1}| + \sup_{\gamma\in[0,1]} |V_{b2}| \\
&= O\left(\frac{N}{n_N}\right) + o_p\left(\frac{N}{n_N}\right) \\
&= o_p\left(\frac{N^2}{n_N}\right).
\end{aligned}$$

Finally, we will show that $\sup_{\gamma\in[0,1]} |V_c| = O\left(\frac{1}{n_N}\right)$. The expression of V_c can be written as

$$\begin{aligned}
V_c &= (1-\gamma)^2 \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{y_i}{\pi_i}\frac{y_j}{\pi_j} \frac{w^*_{sii}}{1-w^*_{sii}} \frac{w^*_{sjj}}{1-w^*_{sjj}} \\
&\quad + (1-\gamma)^2 \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{y_i}{\pi_i}\frac{y_j}{\pi_j} \frac{w^*_{sii}}{1-w^*_{sii}} \frac{w^*_{sjj}}{1-w^*_{sjj}} \left(\frac{I_{ij}}{\pi_{ij}} - 1\right) \\
&= V_{c1} + V_{c2}.
\end{aligned}$$

And,

$$\begin{aligned}
V_{c1} &= (1-\gamma)^2 \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{y_i}{\pi_i}\frac{y_j}{\pi_j} w^*_{sii} w^*_{sjj} \\
&\quad + (1-\gamma)^2 \sum_{i\in U}\sum_{j\in U} \Delta_{ij} \frac{y_i}{\pi_i}\frac{y_j}{\pi_j} w^*_{sii} w^*_{sjj} g^*_{ij}\left(\widehat{\boldsymbol{T}}\right) \\
&= V_{c11} + V_{c12}.
\end{aligned}$$

Suppose assumptions A1, A2 and A4 hold, by Lemma A.2.5, it can be shown that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{c11}| &\leq \frac{1}{\lambda^2} \sum_{i\in U} \left|\pi_i (1-\pi_i) y_i^2\right| \left\{\max_{i\in U} |w^*_{sii}|\right\}^2 \\
&\quad + \frac{1}{\lambda^2} \sum_{(i,j)\in D_{2,N}} \max_{(i,j)\in D_{2,N}} |\Delta_{ij}|\,|y_i y_j| \left\{\max_{i\in U} |w^*_{sii}|\right\}^2 \\
&= O\left(\frac{1}{N}\right) + O\left(\frac{1}{n_N}\right) \\
&= O\left(\frac{1}{n_N}\right),
\end{aligned}$$

and by Lemma A.2.6,

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{c12}| &\leq \frac{1}{\lambda^2}\sum_{i\in U}\left|\pi_i\left(1-\pi_i\right)y_i^2\right|\left\{\max_{i\in U}|w_{sii}^*|\right\}^2 \max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right| \\
&\quad + \frac{1}{\lambda^2}\sum_{(i,j)\in D_{2,N}} \max_{(i,j)\in D_{2,N}} |\Delta_{ij}|\,|y_i y_j|\left\{\max_{i\in U}|w_{sii}^*|\right\}^2 \max_{i,j\in U}\left|g_{ij}^*\left(\widehat{\boldsymbol{T}}\right)\right| \\
&= O\left(\frac{1}{N^2}\right) + O\left(\frac{1}{n_N N}\right) \\
&= O\left(\frac{1}{n_N N}\right).
\end{aligned}$$

Then, we can show that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{c1}| &\leq \sup_{\gamma\in[0,1]} |V_{c11}| + \sup_{\gamma\in[0,1]} |V_{c12}| \\
&= O\left(\frac{1}{n_N}\right) + O\left(\frac{1}{n_N N}\right) \\
&= O\left(\frac{1}{n_N}\right).
\end{aligned}$$

Similarly, by (3.8) and (3.9), V_{c2} has the same order as V_{c1}, which implies that

$$\begin{aligned}
\sup_{\gamma\in[0,1]} |V_{c}| &\leq \sup_{\gamma\in[0,1]} |V_{c1}| + \sup_{\gamma\in[0,1]} |V_{c2}| \\
&= O\left(\frac{1}{n_N}\right) + O\left(\frac{1}{n_N}\right) \\
&= O\left(\frac{1}{n_N}\right).
\end{aligned}$$

Thereby,

$$\begin{aligned}
\sup_{\gamma\in[0,1]}\left|\widehat{V}_{\mathrm{CV}}\left(\bar{y}_\gamma\right) - \frac{1}{N^2}\sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\right| &\leq \frac{1}{N^2}\sup_{\gamma\in[0,1]}\left|V_a - \sum_{i\in U}\sum_{j\in U}\Delta_{ij}\frac{y_i-\gamma y_i^0}{\pi_i}\frac{y_j-\gamma y_j^0}{\pi_j}\right| \\
&\quad + \frac{2}{N^2}\sup_{\gamma\in[0,1]}|V_b| + \frac{1}{N^2}\sup_{\gamma\in[0,1]}|V_c| \\
&= o_p\left(\frac{1}{n_N}\right) + o_p\left(\frac{1}{n_N}\right) + O\left(\frac{1}{n_N N^2}\right) \\
&= o_p\left(\frac{1}{n_N}\right).
\end{aligned}$$

So, the result follows.

□

Corollary 3.2.1. *Let assumptions A1-A6 hold. Then,*

$$\sup_{\gamma\in[0,1]}\left|\widehat{V}_{\mathrm{CV}}\left(\bar{y}_\gamma\right) - \mathrm{MSE}_p\left(\bar{y}_\gamma\right)\right| = o_p\left(n_N^{-1}\right).$$

Thus, based on the theorems derived above, it is possible to use $\widehat{V}_{\mathrm{CV}}(\bar{y}_\gamma)$ as an asymptotically equivalent criterion to $\mathrm{MSE}_p(\bar{y}_\gamma)$ for selecting an optimal parameter γ_{opt}. In section 3.3, we will show the efficiency of this selection criterion.

3.3 Simulation Results

A random population of $N = 1000$ by J values of x is generated from the uniform distribution on $[0,1]$, and 1000 values for the errors ε are drawn from $\mathcal{N}(0,1)$. This one error population is used for all simulations, up to multiplication by σ. Four populations of y are generated as follows:

$$y_{il} = m_l(x_i) + \varepsilon_i \qquad 1 \le i \le 1000, \qquad 1 \le l \le 4,$$

where $\{m_l\}_{l=1}^4$ are predefined functions given in the Table 3.1. Linear model in the regression estimator is used for all functions. The first three functions are linear functions with different dimensions. The fourth function is a quadratic function. The finite population quantities of interest are $\bar{y}_N = \frac{1}{N}\sum_{i=1}^{1000} y_{il}$ for each l.

Population	Expression
1.Linear1	$m_1(\boldsymbol{x}) = 1 + (.5, 1, 2, 1.5, 2.5, 1.2, 2.2, 1)\,\boldsymbol{x}$
2.Linear2	$m_2(\boldsymbol{x}) = 1 + (.5, 1, 2)\,\boldsymbol{x}$
3.Linear3	$m_3(x) = 2x$
4.Quadratic	$m_4(x) = 1 + 2(x - 0.5)^2$

Table 3.1 Four population mean functions.

The samples are drawn by one of two designs, simple random sampling without replacement (SI) or stratified simple random sampling without replacement (STSI). For each simulation run, $M = 1000$ samples are drawn from $\{(x_i, y_i)\}$. For each sample, we compute the estimator $\bar{y}_\gamma$ in equation (3.1) with γ_{opt} and $\widehat{\gamma}_{\mathrm{CV}}$ respectively. Referring to Opsomer and Miller (2005), the optimal mixing parameters γ_{opt} for each population are not sample-based. We compute them by minimizing a simulation-based approximation to the function $\mathrm{MSE}_p(\gamma)$, which is constructed by simulating repeated samples from these populations for a grid of mixing parameters over the interval $[0,1]$, and finding the functions $\mathrm{MSE}_p(\gamma)$ by averaging over these simulations.

For each sample, the mixing parameter $\widehat{\gamma}_{\mathrm{CV}}$ is sought through a search algorithm implemented in R, which uses expression (3.7). A simulation run is determined by sample size n, correlation coefficient R^2 between $\boldsymbol{x}$ and y. For the design of simple random sampling without replacement, simulations are done for $n \in \{100, 200, 500\}$, $R^2 \in [0,1]$. The design of stratified simple random sampling without replacement uses 4 strata with each stratum containing 250 elements, and the stratification of the strata is based on a random variable z_i and ratio r. First, we generate v_i from a standard normal distribution $\mathcal{N}\left(0,\sigma_v^2\right)$ with σ_v^2 satisfying $r = \frac{\sigma^2}{\sigma^2+\sigma_v^2}$. Then z_i's are derived as follows:

$$z_i = \begin{cases} v_i + \varepsilon_i & \text{if } 0 < r < 1 \\ v_i & \text{if } r = 0, \text{ where } \sigma_v^2 = 1 \\ \varepsilon_i & \text{if } r = 1 \end{cases}.$$

After sorting by z_i $(i = 1, \ldots, N)$, the population is separated into 4 strata with boundaries given by equally-spaced quantiles of z. Then, simulations are conducted with the stratum sample sizes $\{(15, 20, 30, 35), (30, 40, 60, 70), (75, 100, 150, 175)\}$, $r \in \{0, 0.25, 0.5, 0.75, 1\}$, $R^2 \in [0, 1]$. Thus, the strata have different sampling rates with the inclusion probability correlated with the model error.

Population	R^2	n	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
1.Linear1	0.010	100	0.000	1414.612	0.109	1430.810	1414.702	1569.084
	0.010	200	0.364	629.399	0.132	631.101	633.276	641.779
	0.010	500	0.455	156.230	0.183	156.794	156.947	157.256
	0.072	100	0.404	192.723	0.334	195.040	197.844	204.280
	0.072	200	0.697	82.702	0.538	83.009	87.363	83.554
	0.072	500	0.960	20.471	0.781	20.545	22.078	20.473
	0.518	100	0.919	14.642	0.912	14.707	29.106	14.748
	0.518	200	0.960	6.020	0.958	6.023	12.405	6.032
	0.518	500	1.000	1.477	0.984	1.480	3.207	1.478
2.Linear2	0.010	100	0.283	397.452	0.109	396.957	398.606	404.542
	0.010	200	0.576	180.755	0.141	181.711	181.916	181.382
	0.010	500	0.525	45.741	0.235	45.865	45.865	45.842
	0.072	100	0.758	52.381	0.444	52.940	55.214	52.668
	0.072	200	1.000	23.613	0.675	23.875	25.293	23.614
	0.072	500	0.889	5.964	0.876	5.970	6.255	5.968
	0.518	100	1.000	3.800	0.956	3.811	7.584	3.802
	0.518	200	1.000	1.704	0.979	1.708	3.388	1.705
	0.518	500	1.000	0.431	0.992	0.431	0.796	0.431
3.Linear3	0.010	100	0.343	295.878	0.103	296.377	296.340	297.637
	0.010	200	0.000	111.933	0.145	112.277	111.934	113.057
	0.010	500	0.253	32.473	0.278	32.534	32.486	32.578
	0.072	100	0.859	38.698	0.592	39.098	40.742	38.750
	0.072	200	0.778	14.661	0.814	14.737	15.349	14.719
	0.072	500	0.828	4.233	0.937	4.239	4.427	4.241
	0.518	100	1.000	2.796	0.981	2.798	5.603	2.797
	0.518	200	0.960	1.061	0.991	1.062	2.264	1.063
	0.518	500	0.960	0.306	0.996	0.306	0.618	0.306
4.Quadratic	0.010	100	0.000	20.598	0.060	20.636	20.598	20.800
	0.010	200	0.000	7.883	0.048	7.907	7.883	7.940
	0.010	500	0.859	2.282	0.040	2.283	2.285	2.282

Continued...

Population	R^2	n	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
	0.072	100	0.000	2.862	0.058	2.876	2.862	2.907
	0.072	200	0.000	1.115	0.044	1.119	1.115	1.127
	0.072	500	0.677	0.320	0.021	0.320	0.320	0.320
	0.518	100	0.000	0.402	0.062	0.407	0.402	0.419
	0.518	200	0.000	0.172	0.055	0.173	0.172	0.176
	0.518	500	0.000	0.046	0.025	0.046	0.046	0.046

Table 3.2: CV mixing parameters $\widehat{\gamma}_{\text{CV}}$ and optimal mixing parameters γ_{opt} with their corresponding MSEs based on 1000 replications of simple random sampling for populations of size $N = 1000$. MSE_0 and MSE_1 are the MSEs for the π estimator and regression estimator respectively.

Table 3.2 shows that the cross-validation mixing parameters $\widehat{\gamma}_{\text{CV}}$ and the optimal mixing parameters γ_{opt} with their corresponding MSEs (N times of real values) for all populations under the design of simple random sampling without replacement. Apparently, in agreement with γ_{opt}, $\widehat{\gamma}_{\text{CV}}$ varies widely across R^2 and between the sample sizes for the three linear populations. Also, CV and optimal mixing parameters are generally an increasing function of the closeness of the relationship between y and x (as measured by R^2). The difference between CV and optimal mixing parameters is generally a decreasing function of the sample size. In the cases where the model is misspecified, corresponding to the cases with quadratic mean for the linear regression estimation, the mixing parameters become very small. $\text{MSE}_p\left(\widehat{\gamma}_{\text{CV}}\right)$ is smaller than $\text{MSE}_p\left(\bar{y}_{\pi}\right)$ for some large R^2 values and smaller than $\text{MSE}_p\left(\bar{y}_{\text{reg}}\right)$ for some small R^2 values.

Population	R^2	n	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
1.Linear1	0.010	100	0.071	842.834	0.103	855.938	843.868	1048.070
	0.010	200	0.141	370.595	0.127	372.221	371.630	407.515
	0.010	500	0.232	108.982	0.183	109.288	109.386	113.610
	0.072	100	0.354	117.671	0.345	120.960	123.332	136.449
	0.072	200	0.535	50.556	0.532	51.178	53.979	53.055
	0.072	500	0.707	14.556	0.749	14.552	15.866	14.791
	0.518	100	0.879	9.590	0.904	9.621	24.569	9.851
	0.518	200	0.949	3.816	0.948	3.798	10.253	3.830
	0.518	500	0.980	1.066	0.971	1.062	2.977	1.068
2.Linear2	0.010	100	0.152	256.449	0.113	258.921	256.949	272.085
	0.010	200	0.182	107.618	0.152	108.122	107.825	111.557
	0.010	500	0.253	29.703	0.241	29.758	29.751	30.127

Continued. . .

Population	R^2	n	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
	0.072	100	0.616	34.539	0.475	35.144	36.859	35.423
	0.072	200	0.677	14.277	0.690	14.368	15.392	14.524
	0.072	500	0.818	3.905	0.848	3.911	4.239	3.922
	0.518	100	0.970	2.553	0.940	2.566	6.509	2.557
	0.518	200	0.949	1.044	0.964	1.041	2.805	1.049
	0.518	500	1.000	0.283	0.977	0.283	0.755	0.283
3.Linear3	0.010	100	0.030	171.780	0.132	173.547	171.787	177.227
	0.010	200	0.515	81.595	0.222	81.839	82.033	81.969
	0.010	500	0.848	24.025	0.414	24.152	24.202	24.031
	0.072	100	0.707	22.768	0.699	23.100	24.514	23.073
	0.072	200	1.000	10.671	0.881	10.718	12.017	10.672
	0.072	500	1.000	3.128	0.946	3.137	3.469	3.129
	0.518	100	0.949	1.659	0.968	1.655	4.890	1.666
	0.518	200	1.000	0.769	0.983	0.771	2.352	0.770
	0.518	500	1.000	0.226	0.991	0.226	0.623	0.226
4.Quadratic	0.010	100	0.000	12.204	0.064	12.252	12.204	12.522
	0.010	200	0.000	5.747	0.063	5.761	5.747	5.828
	0.010	500	0.091	1.674	0.059	1.674	1.674	1.679
	0.072	100	0.000	1.809	0.064	1.822	1.810	1.858
	0.072	200	0.000	0.853	0.058	0.856	0.853	0.863
	0.072	500	0.000	0.240	0.034	0.240	0.240	0.241
	0.518	100	0.000	0.363	0.096	0.368	0.363	0.373
	0.518	200	0.162	0.161	0.092	0.162	0.162	0.163
	0.518	500	0.000	0.043	0.091	0.043	0.043	0.043

Table 3.3: CV mixing parameters $\widehat{\gamma}_{\text{CV}}$ and optimal mixing parameters γ_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling for populations of size $N = 1000$ with $r = 0.5$. MSE_0 and MSE_1 are the MSEs for the π estimator and regression estimator respectively.

The same overall behavior can be seen in the results of Table 3.3, which displays the CV and optimal mixing parameters with their corresponding MSEs for all populations under the design of stratified simple random sampling without replacement with $r = 0.5$.

Population	r	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
1.Linear1	0	0.687	92.947	0.446	94.741	98.491	94.071
	0.25	0.606	82.953	0.468	83.476	86.970	84.614
	0.5	0.535	50.556	0.532	51.178	53.979	53.055

Continued. . .

Population	r	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
	0.75	0.364	36.455	0.628	37.377	38.256	41.905
	1	0.424	15.525	0.635	15.927	18.298	20.510
2.Linear2	0	0.848	25.318	0.744	25.558	26.845	25.364
	0.25	0.667	18.370	0.648	18.623	19.344	18.606
	0.5	0.677	14.277	0.690	14.368	15.392	14.524
	0.75	0.657	8.951	0.759	8.965	9.979	9.238
	1	0.687	4.030	0.803	3.994	5.206	4.265
3.Linear3	0	0.859	17.786	0.721	17.866	18.778	17.812
	0.25	1.000	14.678	0.850	14.790	16.028	14.679
	0.5	1.000	10.671	0.881	10.718	12.017	10.672
	0.75	1.000	7.278	0.912	7.319	9.065	7.279
	1	1.000	2.871	0.961	2.855	4.086	2.872
4.Quadratic	0	0.566	1.337	0.054	1.338	1.339	1.338
	0.25	0.000	1.102	0.048	1.105	1.102	1.118
	0.5	0.000	0.853	0.058	0.856	0.853	0.863
	0.75	0.000	0.616	0.061	0.617	0.616	0.629
	1	0.000	0.293	0.072	0.293	0.293	0.302

Table 3.4: CV mixing parameters $\widehat{\gamma}_{\text{CV}}$ and optimal mixing parameters γ_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling for populations of size $N = 1000$ with $R^2 = 0.072$, $n = (30, 40, 60, 70)$. MSE_0 and MSE_1 are the MSEs for the π estimator and regression estimator respectively.

Table 3.4 displays the CV and optimal mixing parameters with their corresponding MSEs under the design of stratified simple random sampling without replacement with model variance $R^2 = 0.072$ and the stratum sample sizes $n = (30, 40, 60, 70)$. We can find that $\widehat{\gamma}_{\text{CV}}$ changes within a short range, and it tracks γ_{opt} well. Because the decrease of r means that the relationship between the inclusion probability and the model error becomes weaker, it is reasonable to find that the MSE decreases as r increases. $\text{MSE}_p\left(\widehat{\gamma}_{\text{CV}}\right)$ is smaller than $\text{MSE}_p\left(\bar{y}_\pi\right)$ when r is small, and smaller than $\text{MSE}_p\left(\bar{y}_{\text{reg}}\right)$ when r is large.

From the above tables, it can be shown that the estimated mixing parameter usually tracks the optimal parameter quite well. The proposed estimator works better than the π estimator when there is a strong linear relationship between $\boldsymbol{y}$ and $\boldsymbol{x}$, and works better than the regression estimator when there is a weak linear relationship between $\boldsymbol{y}$ and $\boldsymbol{x}$.

3.4 Conclusion

In this chapter, we proposed a new estimator and a design-based CV criterion for selecting the mixing parameter in the estimator. First, we developed theoretical results by proving that the design-based properties of the proposed estimator hold uniformly for a range of mixing parameter values, making mixing parameter selection possible. Then those results were applied to show that the proposed method for mixing parameter selection is asymptotically equivalent to minimizing the MSE of the estimator. By a simulation study, we showed that the estimated mixing parameter usually tracks the optimal parameter quite well. Hence, we recommend the design-based CV criterion whenever data-driven mixing parameter selection for survey estimation is required, except that alternative methods are developed for mixing parameter selection in the finite population estimation context. And the new estimator is recommended instead of using $\bar{y}_\pi$ or $\bar{y}_{\text{reg}}$ alone for estimation.

CHAPTER 4. Model Selection for Regression Estimators

4.1 Definition of the Estimator

In linear regression estimator, how to choose the set of control variables x is a difficult practical problem. Assume that there are J variables of x, then let $\boldsymbol{c} = (c_1, \ldots, c_J)^T$ represent a vector of length J that can only take on values of 0 and 1 in each element. So the number of possible values for the vector is 2^J, which is a finite but potentially very large number. The goal is to select a vector c that minimizes the design MSE of the estimator. Here we will introduce a CV criterion for choosing between combinations of the x variables to be included in the model.

Let $\boldsymbol{X}_c$ represent the matrix with rows $\boldsymbol{x}_{ci}^T$ derived by deleting the element of $\boldsymbol{x}_i^T$ if the corresponding element in $\boldsymbol{c}$ is 0 , for $i \in U$, then the CV variance estimator for

$$\bar{y}_{c,\text{reg}} = \bar{y}_\pi + \left(\bar{\boldsymbol{X}}_{cN} - \bar{\boldsymbol{X}}_{c\pi}\right)^T \widehat{\boldsymbol{\beta}}_c \tag{4.1}$$

is given by

$$\widehat{V}_{\text{CV}}\left(\bar{y}_{c,\text{reg}}\right) = \frac{1}{N^2} \sum_{i,j \in s} \check{\Delta}_{ij} \frac{e_{cis}}{\pi_i} \frac{e_{cjs}}{\pi_j} \frac{1}{1 - w_{csii}^*} \frac{1}{1 - w_{csjj}^*}, \tag{4.2}$$

where $\bar{\boldsymbol{X}}_{cN} = (1/N) \sum_U \boldsymbol{x}_{ci}$, $\bar{\boldsymbol{X}}_{c\pi} = (1/N) \sum_s \boldsymbol{x}_{ci}/\pi_i$, $e_{cis} = y_i - \boldsymbol{x}_{ci}^T \widehat{\boldsymbol{\beta}}_c$, $w_{csii}^* = \boldsymbol{x}_{ci}^T \widehat{\boldsymbol{T}}_c^{-1} \boldsymbol{x}_{ci}/\sigma_i^2 \pi_i$, and $\widehat{\boldsymbol{\beta}}_c$ is given by

$$\begin{aligned} \widehat{\boldsymbol{\beta}}_c &= \widehat{\boldsymbol{T}}_c^{-1} \widehat{\boldsymbol{t}}_c \\ &= \left(\sum_s \frac{\boldsymbol{x}_{ci} \boldsymbol{x}_{ci}^T}{\sigma_i^2 \pi_i}\right)^{-1} \sum_s \frac{\boldsymbol{x}_{ci} y_i}{\sigma_i^2 \pi_i}. \end{aligned} \tag{4.3}$$

Our goal is to use the CV criterion (4.2) as a way to select between the different possible "working models" for the relationship of the y_i and $\boldsymbol{x}_i$ (or equivalently, between different values for the vector $\boldsymbol{c}$). In order to do so, we need our criterion to be consistent for the design MSE of the regression estimator for any value of $\boldsymbol{c}$.

4.2 Theoretical Properties

Similarly, we make the following technical assumptions in order to prove our theoretical results. For simplicity, we will only consider the case with the sample size, denoted by n_N, fixed for each N, and also assume that $n_N \to \infty$. As above, J is fixed.

A1. (Sampling rate $n_N N^{-1}$). As $N \to \infty$, $n_N N^{-1} \to \pi \in (0, 1)$.

A2. (Inclusion probabilities π_i and π_j). For all N, $\min_{i \in U} \pi_i \geq \lambda > 0$, $\min_{i,j \in U} \pi_{ij} \geq \lambda^* > 0$, $\lim_{N \to \infty} n_N \max_{i,j \in U, i \neq j} |\Delta_{ij}| < \infty$.

A3. Assume that $\left(\frac{n_N}{N^2}\widehat{\boldsymbol{T}}_c\right)^{-1}$ and $\left(\frac{n_N}{N^2}\boldsymbol{T}_c\right)^{-1}$ exist for all samples, where $\boldsymbol{T}_c = \sum_U \frac{\boldsymbol{x}_{ci}\boldsymbol{x}_{ci}^T}{\sigma_i^2}$.

A4. $\max_{i \in U} |y_i| < C_y$, and $\max_{i \in U, j \in \{1,\ldots,J\}} |x_{ij}| < C_x$, where C_y and C_x are some positive constants.

A5. $0 < \sigma_L \leq \min_{i \in U} \sigma_i \leq \max_{i \in U} \sigma_i \leq \sigma_U < \infty$, where σ_L and σ_U are some positive constants.

A6. Additional assumptions involving higher-order inclusion probabilities:

$$\lim_{N \to \infty} \max_{(i,j,k) \in D_{3,N}} \left|\mathrm{E}_p\left[(I_i - \pi_i)(I_j I_k - \pi_{jk})\right]\right| = 0$$

$$\lim_{N \to \infty} \max_{(i,j,k,l) \in D_{4,N}} \left|\mathrm{E}_p\left[(I_i I_j - \pi_{ij})(I_k I_l - \pi_{kl})\right]\right| = 0$$

$$\lim_{N \to \infty} n_N \max_{(i,j,k) \in D_{3,N}} \left|\mathrm{E}_p\left[(I_i - \pi_i)^2 (I_j - \pi_j)(I_k - \pi_k)\right]\right| < \infty$$

$$\lim_{N \to \infty} n_N^2 \max_{(i,j,k,l) \in D_{4,N}} \left|\mathrm{E}_p\left[(I_i - \pi_i)(I_j - \pi_j)(I_k - \pi_k)(I_l - \pi_l)\right]\right| < \infty$$

where $D_{t,N}$ denotes the set of all distinct t-tuples from U.

We will therefore establish the following theoretical property:

$$\sup_{\boldsymbol{c}} \left|\widehat{V}_{\mathrm{CV}}\left(\bar{y}_{c,\mathrm{reg}}\right) - \mathrm{MSE}_p\left(\bar{y}_{c,\mathrm{reg}}\right)\right| = o_p\left(n_N^{-1}\right). \tag{4.4}$$

Because there are only a finite and fixed number of possible values for $\boldsymbol{c}$, this property will follow directly from the point-wise (for a fixed $\boldsymbol{c}$) consistency of $\widehat{V}_{\mathrm{CV}}$ for MSE_p. And from the property above, $\widehat{V}_{\mathrm{CV}}\left(\bar{y}_{c,\mathrm{reg}}\right)$ can be used as an asymptotically equivalent criterion to $\mathrm{MSE}_p\left(\bar{y}_{c,\mathrm{reg}}\right)$ for selecting an optimal vector $\boldsymbol{c}_{\mathrm{opt}}$. A simulation study will be given to show the efficiency of this criterion.

4.3 Simulation Results

A random population of $N = 1000$ values of x is generated from the uniform distribution on $[0,1]$, and 1000 values for the errors ε are drawn from $\mathcal{N}(0,1)$. This one error population is used for all simulations, up to multiplication by σ. Twenty populations of y are generated as follows:

$$y_{il} = m_l(x_i) + \varepsilon_i \qquad 1 \le i \le 1000, \qquad 1 \le l \le 20,$$

where $\{m_l\}_{l=1}^{2}0$ are predefined functions given in the Table 4.1. In mean functions m_1 to m_4, a 1000 by 8 matrix of x is generated from a uniform distribution on $[0,1]$ with independent column vectors. Mean functions m_5 to m_{16} are similar to previous functions except that the column vectors are correlated with correlation specified in Table 4.1. To generate a correlated uniform distribution, first we simulate a population of N by J values of matrix from a multivariate normal distribution with specified correlation structure, and then we calculate the cumulative distribution function (cdf) value for each simulated value based on a univariate normal distribution. These generated values follow a multivariate uniform distribution. The correlation structures specified in Table 4.1 for population functions m_5 to m_{16} are used in underlying multivariate normal distribution, not those of multivariate uniform distribution. The main goal is to investigate no correlation, low, medium and high, not the exact numbers. In Mean function m_{17} to m_{19}, the population is defined as a constant β_g. Then $\boldsymbol{x}$ is a vector with gth entry equal to 1 and 0 otherwise, for observation i in group g. For simplicity, all mean functions m_{17} and m_{19} have equal group sizes. In mean function m_{20}, a 1000 by 3 matrix of x is generated from a uniform distribution on $[0,1]$ with independent column vectors.

For post-stratification cases, we do selection from 4 group mean models with 1, 2, 4, and 8 groups. The full regression estimator would always use 8 groups. Then vector $\boldsymbol{c}$ has length 4 with the element equal to 1 for true group mean model, and 0 otherwise.

The finite population quantities of interest are $\bar{y}_N = \frac{1}{N}\sum_{i=1}^{1000} y_{il}$ for each l. A linear regression model is conducted for all population functions, and a quadratic regression model including interaction terms is conducted for quadratic population function compared to a linear regression model.

The samples are drawn by one of two designs, simple random sampling without replacement (SI) or stratified simple random sampling without replacement (STSI). For each simulation run, $M = 1000$ samples are drawn from $\{(x_i, y_i)\}$. For each sample, we compute the estimator $\bar{y}_{c,\text{reg}}$ for

Population	Expression
1.Linear1	$m_1(\boldsymbol{x}) = 1 + (0,1,0,0,0,0,0,0)\,\boldsymbol{x}$
2.Linear2	$m_2(\boldsymbol{x}) = 1 + (0,1,0,1.5,0,0,0,1)\,\boldsymbol{x}$
3.Linear3	$m_3(\boldsymbol{x}) = 1 + (.5,1,0,1.5,2.5,0,0,1)\,\boldsymbol{x}$
4.Linear4	$m_4(\boldsymbol{x}) = 1 + (.5,1,.8,1.5,2.5,2,1,1)\,\boldsymbol{x}$
5.Collinear11	$m_5(\boldsymbol{x}) = 1 + (0,1,0,0,0,0,0,0)\,\boldsymbol{x}$, $\rho = 0.2$ for underlying normal distribution
6.Collinear12	$m_6(\boldsymbol{x}) = 1 + (0,1,0,1.5,0,0,0,1)\,\boldsymbol{x}$, $\rho = 0.2$ for underlying normal distribution
7.Collinear13	$m_7(\boldsymbol{x}) = 1 + (.5,1,0,1.5,2.5,0,0,1)\,\boldsymbol{x}$, $\rho = 0.2$ for underlying normal distribution
8.Collinear14	$m_8(\boldsymbol{x}) = 1 + (.5,1,.8,1.5,2.5,2,1,1)\,\boldsymbol{x}$, $\rho = 0.2$ for underlying normal distribution
9.Collinear21	$m_9(\boldsymbol{x}) = 1 + (0,1,0,0,0,0,0,0)\,\boldsymbol{x}$, $\rho = 0.5$ for underlying normal distribution
10.Collinear22	$m_{10}(\boldsymbol{x}) = 1 + (0,1,0,1.5,0,0,0,1)\,\boldsymbol{x}$, $\rho = 0.5$ for underlying normal distribution
11.Collinear23	$m_{11}(\boldsymbol{x}) = 1 + (.5,1,0,1.5,2.5,0,0,1)\,\boldsymbol{x}$, $\rho = 0.5$ for underlying normal distribution
12.Collinear24	$m_{12}(\boldsymbol{x}) = 1 + (.5,1,.8,1.5,2.5,2,1,1)\,\boldsymbol{x}$, $\rho = 0.5$ for underlying normal distribution
13.Collinear31	$m_{13}(\boldsymbol{x}) = 1 + (0,1,0,0,0,0,0,0)\,\boldsymbol{x}$, $\rho = 0.98$ for underlying normal distribution
14.Collinear32	$m_{14}(\boldsymbol{x}) = 1 + (0,1,0,1.5,0,0,0,1)\,\boldsymbol{x}$, $\rho = 0.98$ for underlying normal distribution
15.Collinear33	$m_{15}(\boldsymbol{x}) = 1 + (.5,1,0,1.5,2.5,0,0,1)\,\boldsymbol{x}$, $\rho = 0.98$ for underlying normal distribution
16.Collinear34	$m_{16}(\boldsymbol{x}) = 1 + (.5,1,.8,1.5,2.5,2,1,1)\,\boldsymbol{x}$, $\rho = 0.98$ for underlying normal distribution
17.Post-stratification1	$m_{17}(x) = 0.5$
18.Post-stratification2	$m_{18}(\boldsymbol{x}) = (0.5,1,1.5,2)\,\boldsymbol{x}$
19.Post-stratification3	$m_{19}(\boldsymbol{x}) = (0.5,1,1.5,2,2.5,3,3.5,4)\,\boldsymbol{x}$
20.Quadratic	$m_{20}(\boldsymbol{x}) = 1 + 2\left((0,1,0)\,\boldsymbol{x} - 0.5\right)^2$

Table 4.1 Population mean functions.

a given $\widehat{\boldsymbol{c}}_{\text{CV}}$, and variance of weights for both CV estimator and full regression model estimator. The vector $\widehat{\boldsymbol{c}}_{\text{CV}}$ is sought by minimizing $\widehat{V}_{\text{CV}}\left(\bar{y}_{c,\text{reg}}\right)$. A simulation run is determined by sample size n, error variance σ^2. For the design of simple random sampling without replacement, simulations are done for $n \in \{100, 200, 500\}$, $\sigma^2 \in \{0.01, 0.16\}$. The design of stratified simple random sampling without replacement uses 4 strata with each stratum containing 250 elements, and the stratification of the strata is based on a random variable z_i and ratio r. First, we generate v_i from a standard normal distribution $\mathcal{N}\left(0, \sigma_v^2\right)$ with σ_v^2 satisfying $r = \frac{\sigma^2}{\sigma^2 + \sigma_v^2}$. Then z_i's are derived as follows:

$$z_i = \begin{cases} v_i + \varepsilon_i & \text{if } 0 < r < 1 \\ v_i & \text{if } r = 0, \text{ where } \sigma_v^2 = 1 \\ \varepsilon_i & \text{if } r = 1 \end{cases}.$$

After sorting by z_i $(i = 1, \ldots, N)$, the population is separated into 4 strata with boundaries given by equally-spaced quantiles of z. Then, simulations are conducted with the stratum sample sizes $\{(15, 20, 30, 35), (30, 40, 60, 70), (75, 100, 150, 175)\}$, $r \in \{0, 0.25, 0.5, 0.75, 1\}$, $\sigma^2 \in \{0.01, 0.16\}$. Thus, the strata have different sampling rates with the inclusion probability correlated with the model error.

Population	σ	n	$\mathrm{E}\left(\widehat{\boldsymbol{c}}_{\text{CV}}\right)$	$\boldsymbol{c}$	δ_{CV}	δ_π	δ_{reg}	R_{w}
1.Linear1	0.1	100	(0.15,1,0.15,0.15,0.13,0.19,0.17,0.17)	(0,1,0,0,0,0,0,0)	1.07	9.28	1.11	0.25
		200	(0.13,1,0.13,0.14,0.12,0.21,0.20,0.15)		1.01	8.25	1.03	0.24
		500	(0.05,1,0.04,0.08,0.04,0.33,0.20,0.09)		1.01	9.05	1.02	0.24
	0.4	100	(0.15,1,0.15,0.15,0.13,0.19,0.17,0.17)		1.07	1.52	1.11	0.25
		200	(0.13,1,0.13,0.14,0.12,0.21,0.20,0.15)		1.01	1.33	1.03	0.24
		500	(0.05,1,0.04,0.08,0.04,0.33,0.20,0.09)		1.01	1.41	1.02	0.24
2.Linear2	0.1	100	(0.15,1,0.15,1,0.13,0.19,0.18,1)	(0,1,0,1,0,0,0,1)	1.05	41.03	1.08	0.47
		200	(0.13,1,0.13,1,0.12,0.20,0.20,1)		1.00	38.46	1.01	0.45
		500	(0.05,1,0.04,1,0.04,0.33,0.20,1)		1.00	38.88	1.01	0.45
	0.4	100	(0.15,1,0.15,1,0.13,0.19,0.18,1)		1.05	3.42	1.08	0.47
		200	(0.13,1,0.13,1,0.12,0.20,0.20,1)		1.00	3.18	1.01	0.45

Continued...

Population	σ	n	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
		500	(0.05,1,0.04,1,0.04,0.33,0.20,1)		1.00	3.31	1.01	0.45
3.Linear3	0.1	100	(1,1,0.16,1,1,0.18,0.17,1)	(1,1,0,1,1,0,0,1)	1.02	98.35	1.03	0.68
		200	(1,1,0.13,1,1,0.20,0.20,1)		1.00	108.68	1.00	0.68
		500	(1,1,0.04,1,1,0.33,0.20,1)		1.00	104.37	1.01	0.70
	0.4	100	(0.98,1,0.16,1,1,0.18,0.17,1)		1.02	6.95	1.03	0.67
		200	(1,1,0.13,1,1,0.20,0.20,1)		1.00	7.57	1.00	0.68
		500	(1,1,0.04,1,1,0.33,0.20,1)		1.00	7.40	1.01	0.70
4.Linear4	0.1	100	(1,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.00	132.30	1.00	1.00
		200	(1,1,1,1,1,1,1,1)		1.00	141.27	1.00	1.00
		500	(1,1,1,1,1,1,1,1)		1.00	141.53	1.00	1.00
	0.4	100	(0.98,1,1,1,1,1,1,1)		1.00	9.18	1.00	1.00
		200	(1,1,1,1,1,1,1,1)		1.00	9.77	1.00	1.00
		500	(1,1,1,1,1,1,1,1)		1.00	9.93	1.00	1.00
5.Collinear11	0.1	100	(0.14,1,0.18,0.27,0.16,0.16,0.14,0.17)	(0,1,0,0,0,0,0,0)	1.03	8.94	1.05	0.26
		200	(0.13,1,0.18,0.42,0.11,0.17,0.10,0.17)		1.02	9.11	1.02	0.28
		500	(0.08,1,0.22,0.75,0.05,0.18,0.06,0.18)		1.01	9.81	1.01	0.32
	0.4	100	(0.14,1,0.18,0.27,0.16,0.16,0.14,0.17)		1.03	1.44	1.05	0.26
		200	(0.13,1,0.18,0.42,0.11,0.17,0.10,0.17)		1.02	1.44	1.02	0.28
		500	(0.08,1,0.22,0.75,0.05,0.18,0.06,0.18)		1.01	1.54	1.01	0.32
6.Collinear12	0.1	100	(0.14,1,0.20,1,0.15,0.16,0.16,1)	(0,1,0,1,0,0,0,1)	1.03	49.88	1.03	0.45
		200	(0.13,1,0.18,1,0.12,0.18,0.10,1)		1.02	53.07	1.02	0.46
		500	(0.08,1,0.22,1,0.05,0.17,0.06,1)		1.00	54.46	1.01	0.45
	0.4	100	(0.14,1,0.20,1,0.15,0.16,0.16,1)		1.03	4.00	1.03	0.45
		200	(0.13,1,0.18,1,0.12,0.18,0.10,1)		1.02	4.27	1.02	0.46
		500	(0.08,1,0.22,1,0.05,0.17,0.06,1)		1.00	4.48	1.01	0.45
7.Collinear13	0.1	100	(1,1,0.21,1,1,0.16,0.14,1)	(1,1,0,1,1,0,0,1)	1.02	139.82	1.01	0.66
		200	(1,1,0.19,1,1,0.17,0.10,1)		1.01	141.51	1.02	0.68
		500	(1,1,0.22,1,1,0.16,0.06,1)		1.01	151.21	1.01	0.68
	0.4	100	(0.96,1,0.2,1,1,0.16,0.14,1)		1.02	9.57	1.01	0.66
		200	(1,1,0.19,1,1,0.17,0.10,1)		1.01	9.85	1.02	0.68
		500	(1,1,0.22,1,1,0.16,0.06,1)		1.01	10.69	1.01	0.68
8.Collinear14	0.1	100	(1,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.00	266.46	1.00	1.00
		200	(1,1,1,1,1,1,1,1)		1.00	271.18	1.00	1.00
		500	(1,1,1,1,1,1,1,1)		1.00	282.15	1.00	1.00
	0.4	100	(0.96,1,1,1,1,1,1,1)		1.00	17.34	1.00	0.99
		200	(1,1,1,1,1,1,1,1)		1.00	17.84	1.00	1.00
		500	(1,1,1,1,1,1,1,1)		1.00	18.68	1.00	1.00
9.Collinear21	0.1	100	(0.12,1,0.16,0.27,0.16,0.16,0.14,0.16)	(0,1,0,0,0,0,0,0)	1.03	9.18	1.04	0.26
		200	(0.12,1,0.17,0.39,0.11,0.18,0.10,0.18)		1.01	9.31	1.02	0.27
		500	(0.09,1,0.22,0.74,0.05,0.26,0.08,0.18)		1.01	10.06	1.01	0.34
	0.4	100	(0.12,1,0.16,0.27,0.16,0.16,0.14,0.16)		1.03	1.46	1.04	0.26
		200	(0.12,1,0.17,0.39,0.11,0.18,0.10,0.18)		1.01	1.46	1.02	0.27
		500	(0.09,1,0.22,0.74,0.05,0.26,0.08,0.18)		1.01	1.56	1.01	0.34
10.Collinear22	0.1	100	(0.14,1,0.20,1,0.15,0.18,0.14,1)	(0,1,0,1,0,0,0,1)	1.02	69.04	1.03	0.45
		200	(0.13,1,0.17,1,0.12,0.20,0.11,1)		1.02	72.85	1.02	0.46
		500	(0.09,1,0.21,1,0.04,0.25,0.08,1)		1.00	76.62	1.01	0.46
	0.4	100	(0.14,1,0.20,1,0.15,0.18,0.14,1)		1.02	5.17	1.03	0.45
		200	(0.13,1,0.17,1,0.12,0.20,0.11,1)		1.02	5.49	1.02	0.46
		500	(0.09,1,0.21,1,0.04,0.25,0.08,1)		1.00	5.84	1.01	0.46
11.Collinear23	0.1	100	(1,1,0.20,1,1,0.17,0.14,1)	(1,1,0,1,1,0,0,1)	1.01	213.71	1.01	0.67
		200	(1,1,0.18,1,1,0.20,0.10,1)		1.01	218.62	1.01	0.68
		500	(1,1,0.21,1,1,0.23,0.08,1)		1.00	232.52	1.00	0.69
	0.4	100	(0.87,1,0.19,1,1,0.17,0.14,1)		1.02	14.17	1.01	0.65
		200	(0.99,1,0.18,1,1,0.20,0.10,1)		1.01	14.62	1.01	0.68
		500	(1,1,0.21,1,1,0.23,0.08,1)		1.00	15.73	1.00	0.69
12.Collinear24	0.1	100	(1,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.00	475.20	1.00	1.00
		200	(1,1,1,1,1,1,1,1)		1.00	486.97	1.00	1.00
		500	(1,1,1,1,1,1,1,1)		1.00	513.69	1.00	1.00
	0.4	100	(0.87,1,0.99,1,1,1,1,1)		1.01	30.34	1.00	0.98
		200	(0.99,1,1,1,1,1,1,1)		1.00	31.27	1.00	1.00
		500	(1,1,1,1,1,1,1,1)		1.00	33.14	1.00	1.00

Continued...

Population	σ	n	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
13.Collinear31	0.1	100	(0.13,1,0.14,0.24,0.16,0.17,0.12,0.14)	(0,1,0,0,0,0,0,0)	1.03	9.27	1.05	0.25
		200	(0.11,1,0.12,0.30,0.14,0.18,0.09,0.13)		1.01	9.27	1.02	0.25
		500	(0.06,1,0.14,0.55,0.13,0.29,0.07,0.14)		1.01	10.03	1.01	0.30
	0.4	100	(0.15,0.39,0.17,0.34,0.23,0.19,0.15,0.17)		1.02	1.47	1.05	0.21
		200	(0.10,0.52,0.13,0.38,0.18,0.17,0.10,0.11)		1.02	1.48	1.02	0.20
		500	(0.04,0.78,0.11,0.55,0.14,0.23,0.06,0.11)		1.00	1.56	1.01	0.26
14.Collinear32	0.1	100	(0.14,1,0.17,1,0.16,0.19,0.12,0.99)	(0,1,0,1,0,0,0,1)	1.03	100.01	1.03	0.45
		200	(0.11,1,0.15,1,0.13,0.22,0.11,1)		1.02	101.79	1.02	0.46
		500	(0.07,1,0.15,1,0.12,0.34,0.08,1)		1.00	109.68	1.00	0.47
	0.4	100	(0.13,0.35,0.15,0.86,0.18,0.14,0.13,0.31)		1.02	7.06	1.03	0.26
		200	(0.10,0.57,0.12,0.99,0.15,0.18,0.12,0.40)		1.02	7.22	1.02	0.32
		500	(0.05,0.87,0.12,1,0.14,0.28,0.07,0.70)		1.00	7.83	1.00	0.40
15.Collinear33	0.1	100	(0.74,1,0.19,1,1,0.20,0.13,1)	(1,1,0,1,1,0,0,1)	1.02	337.45	1.02	0.64
		200	(0.96,1,0.17,1,1,0.22,0.11,1)		1.01	342.89	1.01	0.68
		500	(1,1,0.17,1,1,0.34,0.07,1)		1.00	369.73	1.00	0.69
	0.4	100	(0.22,0.39,0.16,0.84,0.95,0.16,0.14,0.32)		1.01	21.82	1.02	0.37
		200	(0.26,0.56,0.14,0.98,1,0.17,0.11,0.42)		1.01	22.23	1.01	0.44
		500	(0.29,0.88,0.12,1,1,0.25,0.07,0.73)		1.00	24.13	1.00	0.54
16.Collinear34	0.1	100	(0.75,1,0.96,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.01	826.56	1.00	0.96
		200	(0.96,1,1,1,1,1,1,1)		1.00	844.18	1.00	1.00
		500	(1,1,1,1,1,1,1,1)		1.00	919.22	1.00	1.00
	0.4	100	(0.27,0.46,0.31,0.85,0.96,0.72,0.46,0.36)		1.02	52.24	1.00	0.52
		200	(0.31,0.62,0.37,0.98,1,0.93,0.64,0.50)		1.02	53.45	1.00	0.65
		500	(0.39,0.91,0.58,1,1,1,0.94,0.82)		1.01	58.45	1.00	0.82
17.Post-ST1	0.1	100	(0.78,0.13,0.06,0.02)	(1,0,0,0)	1.02	1.00	1.09	0.08
		200	(0.82,0.10,0.05,0.03)		1.01	1.00	1.03	0.07
		500	(0.85,0.10,0.04,0.02)		1.00	1.00	1.01	0.05
	0.4	100	(0.78,0.13,0.06,0.02)		1.02	1.00	1.09	0.08
		200	(0.82,0.10,0.05,0.03)		1.01	1.00	1.03	0.07
		500	(0.85,0.10,0.04,0.02)		1.00	1.00	1.01	0.05
18.Post-ST2	0.1	100	(0,0,0.90,0.10)	(0,0,1,0)	1.02	31.69	1.05	0.45
		200	(0,0,0.90,0.10)		1.01	38.74	1.02	0.48
		500	(0,0,0.91,0.09)		1.00	33.56	1.01	0.48
	0.4	100	(0,0,0.90,0.10)		1.02	2.84	1.05	0.45
		200	(0,0,0.90,0.10)		1.01	3.21	1.02	0.48
		500	(0,0,0.91,0.09)		1.00	2.89	1.01	0.48
19.Post-ST3	0.1	100	(0,0,0,1)	(0,0,0,1)	1.00	123.68	1.00	1.00
		200	(0,0,0,1)		1.00	156.23	1.00	1.00
		500	(0,0,0,1)		1.00	140.32	1.00	1.00
	0.4	100	(0,0,0,1)		1.00	8.47	1.00	1.00
		200	(0,0,0,1)		1.00	10.32	1.00	1.00
		500	(0,0,0,1)		1.00	9.39	1.00	1.00
20.Quadratic1	0.1	100	(0.15,0.17,0.14)	(0,0,0)	1.04	1.00	1.05	0.20
		200	(0.13,0.15,0.11)	(0,0,0)	1.02	1.00	1.02	0.16
		500	(0.08,0.14,0.08)	(0,0,0)	1.01	1.00	1.01	0.13
	0.4	100	(0.14,0.14,0.16)	(0,0,0)	1.01	1.00	1.03	0.15
		200	(0.10,0.12,0.14)	(1,0,0)	1.01	1.00	1.01	0.11
		500	(0.05,0.06,0.12)	(0,0,0)	1.00	1.00	1.00	0.07
21.Quadratic2	0.1	100	(0.09,0.77,0.10,0.08,0.77,0.09,0.11,0.09,0.20)	(0,1,0,0,1,0,0,0,0)	1.44	3.05	1.10	0.24
		200	(0.04,0.96,0.05,0.08,0.96,0.07,0.11,0.09,0.18)	(0,1,0,1,1,0,0,0,0)	1.13	2.81	1.05	0.25
		500	(0.02,1,0.04,0.03,1,0.07,0.05,0.04,0.13)	(0,1,0,0,1,0,0,0,0)	1.00	2.81	1.01	0.26
	0.4	100	(0.08,0.31,0.08,0.08,0.32,0.08,0.06,0.08,0.14)	(0,1,0,0,1,0,0,0,0)	1.11	1.11	1.10	0.12
		200	(0.06,0.44,0.06,0.08,0.44,0.06,0.07,0.08,0.14)	(0,1,0,1,1,0,0,0,0)	1.06	1.10	1.05	0.14
		500	(0.02,0.76,0.04,0.04,0.76,0.06,0.05,0.05,0.16)	(0,1,0,0,1,0,0,0,0)	1.02	1.09	1.01	0.21

Continued...

Population	σ	n	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}

Table 4.2: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$, true or optimal $\boldsymbol{c}$, ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true (for linear, collinear and post-stratification functions) or optimal (for quadratic functions) estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of simple random sampling based on linear regression model for linear1-quadratic1, and quadratic model for quadratic2.

Table 4.2 shows the expected vector of $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$, true (for population 1-19) or optimal (for population 20-21) $\boldsymbol{c}$, MSE ratios ($\delta = \frac{1}{M}\sum_{i=1}^{M}\left(\widehat{\bar{y}}_i - \bar{y}_N\right)^2 / \frac{1}{M}\sum_{i=1}^{M}\left(\widehat{\bar{y}}_{\mathrm{tr/opt},i} - \bar{y}_N\right)^2$), and ratios of variance of weights $R_w = \frac{1}{M}\sum_{i=1}^{M}\mathrm{Var}\left(g_{\mathrm{CV},ks,i}/\pi_{k,i}\right) / \frac{1}{M}\sum_{i=1}^{M}\mathrm{Var}\left(g_{\mathrm{reg},ks,i}/\pi_{k,i}\right)$ under the design of SI with $n \in \{100, 200, 500\}$ and $\sigma^2 \in \{0.01, 0.16\}$. For quadratic1 population, since there is no true vector of $\boldsymbol{c}$ for a linear regression model, an optimal $\boldsymbol{c}$ vector and its corresponding MSE will be computed by finding the minimal MSE based on the 2^J possible vectors of $\boldsymbol{c}$. In our case, $J = 3$ is introduced. Also a quadratic regression model is fitted for quadratic2 population function with the elements of $\boldsymbol{c}$ corresponding to variables: x_1, x_2, x_3, x_1^2, x_2^2, x_3^2, x_1x_2, x_1x_3 and x_2x_3. Apparently, there is little difference between CV and full regression estimators for all population functions in terms of MSE ratios. Both CV and full regression estimators have much smaller MSEs than π estimator for most of population functions except for post-stratification1 and quadratic1 functions, especially when standard deviation σ is small. Also both MSE ratios of CV and full regression estimators are stable across standard deviation σ, sample size, and dimension of variable $\boldsymbol{x}$, while the MSE ratio of π estimator is obviously a decreasing function of the closeness of the relationship between y and x (as measured by σ^2), and an increasing function of dimension of variable $\boldsymbol{x}$. By checking the expected vector of $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$, we know that for high dimensional collinear cases with high correlation, CV estimator is as efficient as full regression and true estimators even with only a subset selected for the model. Obviously, CV estimator has less various weights than full regression estimator. The ratio of variance of weights increases as the dimension of $\boldsymbol{x}$ increases in linear and collinear cases, and it is slightly an increasing function of sample size for all functions except post-stratification1 and quadratic1 functions. Also it is generally stable across standard deviation except for high correlated collinear cases with high dimension.

Population	σ	n	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
1.Linear1	0.1	100	(0.15,1,0.15,0.15,0.16,0.18,0.15,0.16)	(0,1,0,0,0,0,0,0)	1.05	15.60	1.22	0.70
		200	(0.12,1,0.20,0.14,0.15,0.21,0.12,0.13)		1.02	16.57	1.12	0.85
		500	(0.10,1,0.25,0.11,0.13,0.24,0.06,0.09)		1.01	16.06	1.05	0.96
	0.4	100	(0.15,1,0.15,0.15,0.16,0.18,0.15,0.16)		1.05	1.82	1.22	0.70

Continued...

Population	σ	n	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
		200	(0.12,1,0.20,0.14,0.15,0.21,0.12,0.13)		1.02	1.83	1.12	0.85
		500	(0.10,1,0.25,0.11,0.13,0.24,0.06,0.09)		1.01	1.84	1.05	0.96
2.Linear2	0.1	100	(0.13,1,0.15,1,0.16,0.18,0.15,1)	(0,1,0,1,0,0,0,1)	1.04	71.48	1.17	0.78
		200	(0.12,1,0.20,1,0.16,0.21,0.12,1)		1.02	76.84	1.08	0.89
		500	(0.10,1,0.24,1,0.13,0.24,0.06,1)		1.01	69.76	1.06	0.97
	0.4	100	(0.13,1,0.15,1,0.16,0.18,0.15,1)		1.04	5.28	1.17	0.78
		200	(0.12,1,0.20,1,0.16,0.21,0.12,1)		1.02	5.47	1.08	0.89
		500	(0.10,1,0.24,1,0.13,0.24,0.06,1)		1.01	5.11	1.06	0.97
3.Linear3	0.1	100	(1,1,0.15,1,1,0.18,0.14,1)	(1,1,0,1,1,0,0,1)	1.02	164.57	1.08	0.87
		200	(1,1,0.21,1,1,0.22,0.12,1)		1.01	191.57	1.04	0.93
		500	(1,1,0.24,1,1,0.24,0.06,1)		1.01	176.65	1.04	0.98
	0.4	100	(0.98,1,0.15,1,1,0.18,0.14,1)		1.02	11.02	1.08	0.87
		200	(1,1,0.21,1,1,0.22,0.12,1)		1.01	12.83	1.04	0.93
		500	(1,1,0.24,1,1,0.24,0.06,1)		1.01	11.87	1.04	0.98
4.Linear4	0.1	100	(1,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.00	215.24	1.00	1.00
		200	(1,1,1,1,1,1,1,1)		1.00	237.12	1.00	1.00
		500	(1,1,1,1,1,1,1,1)		1.00	244.51	1.00	1.00
	0.4	100	(0.98,1,1,1,1,1,1,1)		1.00	14.13	1.00	1.00
		200	(1,1,1,1,1,1,1,1)		1.00	15.75	1.00	1.00
		500	(1,1,1,1,1,1,1,1)		1.00	16.01	1.00	1.00
5.Collinear11	0.1	100	(0.42,1,0.15,0.26,0.15,0.14,0.15,0.15)	(0,1,0,0,0,0,0,0)	1.01	14.44	1.19	0.72
		200	(0.49,1,0.16,0.38,0.16,0.15,0.14,0.14)		0.96	16.06	1.05	0.85
		500	(0.62,1,0.17,0.68,0.18,0.17,0.13,0.08)		0.96	16.12	0.99	0.97
	0.4	100	(0.42,1,0.15,0.26,0.15,0.14,0.15,0.15)		1.01	1.72	1.19	0.72
		200	(0.49,1,0.16,0.38,0.16,0.15,0.14,0.14)		0.96	1.81	1.05	0.85
		500	(0.62,1,0.17,0.68,0.18,0.17,0.13,0.08)		0.96	1.75	0.99	0.97
6.Collinear12	0.1	100	(0.45,1,0.14,1,0.14,0.14,0.16,1)	(0,1,0,1,0,0,0,1)	0.98	78.97	1.13	0.79
		200	(0.53,1,0.16,1,0.15,0.16,0.13,1)		0.97	98.49	1.05	0.90
		500	(0.66,1,0.17,1,0.18,0.17,0.13,1)		0.97	103.22	0.99	0.98
	0.4	100	(0.45,1,0.14,1,0.14,0.14,0.16,1)		0.98	5.62	1.13	0.79
		200	(0.53,1,0.16,1,0.15,0.16,0.13,1)		0.97	7.07	1.05	0.90
		500	(0.66,1,0.17,1,0.18,0.17,0.13,1)		0.97	7.23	0.99	0.98
7.Collinear13	0.1	100	(0.43,1,0.17,1,1,0.23,0.16,1)	(1,1,0,1,1,0,0,1)	1.05	208.88	1.10	0.83
		200	(0.28,1,0.26,1,1,0.39,0.20,1)		1.12	261.74	1.04	0.91
		500	(0.10,1,0.47,1,1,0.80,0.34,1)		1.07	296.16	1.02	0.97
	0.4	100	(0,1,0.13,1,1,0.11,0.13,1)		0.69	13.13	1.10	0.79
		200	(0,1,0.13,1,1,0.12,0.10,1)		0.67	16.89	1.04	0.88
		500	(0,1,0.14,1,1,0.11,0.10,1)		0.66	18.80	1.02	0.96
8.Collinear14	0.1	100	(0.60,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.06	382.33	1.00	0.97
		200	(0.35,1,1,1,1,1,1,1)		1.15	490.75	1.00	0.96
		500	(0.08,1,1,1,1,1,1,1)		1.11	561.78	1.00	0.98
	0.4	100	(0,1,1,1,1,1,1,1)		0.73	23.74	1.00	0.92
		200	(0,1,1,1,1,1,1,1)		0.69	30.88	1.00	0.95
		500	(0,1,1,1,1,1,1,1)		0.67	34.89	1.00	0.98
9.Collinear21	0.1	100	(0.44,1,0.13,0.29,0.16,0.12,0.15,0.15)	(0,1,0,0,0,0,0,0)	0.99	14.04	1.20	0.72
		200	(0.50,1,0.16,0.41,0.17,0.15,0.14,0.15)		0.95	15.90	1.05	0.86
		500	(0.62,1,0.14,0.68,0.20,0.17,0.12,0.07)		0.96	16.59	1.00	0.97
	0.4	100	(0.44,1,0.13,0.29,0.16,0.12,0.15,0.15)		0.99	1.63	1.20	0.72
		200	(0.50,1,0.16,0.41,0.17,0.15,0.14,0.15)		0.95	1.74	1.05	0.86
		500	(0.62,1,0.14,0.68,0.20,0.17,0.12,0.07)		0.96	1.72	1.00	0.97
10.Collinear22	0.1	100	(0.46,1,0.14,1,0.16,0.14,0.15,1)	(0,1,0,1,0,0,0,1)	0.98	107.45	1.14	0.79
		200	(0.56,1,0.15,1,0.17,0.16,0.15,1)		0.97	132.26	1.05	0.90
		500	(0.68,1,0.13,1,0.21,0.16,0.13,1)		0.96	140.38	0.99	0.98
	0.4	100	(0.46,1,0.14,1,0.15,0.14,0.15,1)		0.98	7.15	1.14	0.79
		200	(0.56,1,0.15,1,0.17,0.16,0.15,1)		0.97	8.92	1.05	0.90
		500	(0.68,1,0.13,1,0.21,0.16,0.13,1)		0.96	9.23	0.99	0.98
11.Collinear23	0.1	100	(0.20,1,0.25,1,1,0.36,0.25,1)	(1,1,0,1,1,0,0,1)	0.99	318.63	1.11	0.82
		200	(0.09,1,0.39,1,1,0.62,0.37,1)		1.02	394.06	1.05	0.91
		500	(0,1,0.68,1,1,0.96,0.68,1)		0.94	444.95	1.02	0.98
	0.4	100	(0,1,0.14,1,1,0.12,0.13,1)		0.75	19.62	1.11	0.79
		200	(0,1,0.12,1,1,0.12,0.11,1)		0.73	24.75	1.05	0.88

Continued...

Population	σ	n	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
		500	(0,1,0.12,1,1,0.11,0.12,1)		0.74	27.58	1.02	0.96
12.Collinear24	0.1	100	(0.29,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	0.99	673.17	1.00	0.94
		200	(0.09,1,1,1,1,1,1,1)		1.01	863.80	1.00	0.95
		500	(0,1,1,1,1,1,1,1)		0.94	984.91	1.00	0.98
	0.4	100	(0,1,1,1,1,1,1,1)		0.77	41.34	1.00	0.92
		200	(0,1,1,1,1,1,1,1)		0.75	53.56	1.00	0.95
		500	(0,1,1,1,1,1,1,1)		0.74	60.55	1.00	0.98
13.Collinear31	0.1	100	(0.48,1,0.13,0.28,0.21,0.12,0.15,0.14)	(0,1,0,0,0,0,0,0)	1.00	13.10	1.22	0.73
		200	(0.55,1,0.13,0.44,0.25,0.11,0.14,0.15)		0.95	15.09	1.05	0.88
		500	(0.70,1,0.08,0.73,0.34,0.11,0.11,0.07)		0.96	16.04	1.01	0.98
	0.4	100	(0.39,0.38,0.16,0.33,0.27,0.13,0.18,0.21)		1.01	1.48	1.22	0.70
		200	(0.37,0.46,0.13,0.45,0.28,0.12,0.15,0.18)		0.97	1.63	1.05	0.86
		500	(0.35,0.45,0.09,0.61,0.30,0.11,0.10,0.12)		0.97	1.64	1.01	0.96
14.Collinear32	0.1	100	(0.44,0.99,0.14,1,0.20,0.14,0.16,1)	(0,1,0,1,0,0,0,1)	0.99	152.46	1.17	0.80
		200	(0.54,1,0.13,1,0.22,0.13,0.16,1)		0.97	181.41	1.05	0.91
		500	(0.66,1,0.10,1,0.33,0.15,0.12,1)		0.97	194.01	1.01	0.98
	0.4	100	(0.26,0.34,0.13,0.80,0.22,0.13,0.18,0.51)		1.01	9.55	1.17	0.72
		200	(0.35,0.47,0.12,0.94,0.24,0.13,0.15,0.70)		0.98	11.50	1.05	0.88
		500	(0.38,0.52,0.10,1,0.29,0.20,0.11,0.96)		0.97	12.08	1.01	0.97
15.Collinear33	0.1	100	(0.02,1,0.13,1,1,0.12,0.16,1)	(1,1,0,1,1,0,0,1)	0.82	501.89	1.12	0.79
		200	(0,1,0.13,1,1,0.12,0.14,1)		0.81	622.08	1.05	0.88
		500	(0,1,0.14,1,1,0.1,0.17,1)		0.82	681.24	1.03	0.96
	0.4	100	(0.14,0.34,0.13,0.81,0.96,0.14,0.19,0.53)		0.95	30.63	1.12	0.75
		200	(0.14,0.41,0.13,0.94,1,0.14,0.14,0.75)		0.94	38.34	1.05	0.88
		500	(0.04,0.38,0.12,1,1,0.27,0.11,0.98)		0.95	41.61	1.03	0.96
16.Collinear34	0.1	100	(0.06,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	0.84	1135.78	1.00	0.93
		200	(0.04,1,1,1,1,1,1,1)		0.82	1491.56	1.00	0.95
		500	(0,1,1,1,1,1,1,1)		0.81	1661.60	1.00	0.98
	0.4	100	(0.10,0.37,0.38,0.82,0.97,0.79,0.55,0.60)		0.92	69.51	1.00	0.81
		200	(0.10,0.46,0.48,0.95,1,0.95,0.72,0.79)		0.94	91.85	1.00	0.92
		500	(0.02,0.45,0.75,1,1,1,0.91,0.98)		0.94	101.78	1.00	0.98
17.Post-ST1	0.1	100	(0.77,0.15,0.06,0.02)	(1,0,0,0)	1.02	1.00	1.19	0.62
		200	(0.75,0.16,0.06,0.03)		1.01	1.00	1.08	0.81
		500	(0.68,0.18,0.07,0.06)		1.01	1.00	1.04	0.95
	0.4	100	(0.77,0.15,0.06,0.02)		1.02	1.00	1.19	0.62
		200	(0.75,0.16,0.06,0.03)		1.01	1.00	1.08	0.81
		500	(0.68,0.18,0.07,0.06)		1.01	1.00	1.04	0.95
18.Post-ST2	0.1	100	(0,0,0.90,0.10)	(0,0,1,0)	1.02	60.74	1.10	0.62
		200	(0,0,0.91,0.09)		1.01	60.37	1.04	0.81
		500	(0,0,0.84,0.16)		1.01	54.45	1.03	0.95
	0.4	100	(0,0,0.90,0.10)		1.02	4.39	1.10	0.62
		200	(0,0,0.91,0.09)		1.01	4.63	1.04	0.81
		500	(0,0,0.84,0.16)		1.01	4.16	1.03	0.95
19.Post-ST3	0.1	100	(0,0,0,1)	(0,0,0,1)	1.00	234.37	1.00	0.62
		200	(0,0,0,1)		1.00	244.96	1.00	0.81
		500	(0,0,0,1)		1.00	220.46	1.00	0.95
	0.4	100	(0,0,0,1)		1.00	14.90	1.00	0.62
		200	(0,0,0,1)		1.00	16.10	1.00	0.81
		500	(0,0,0,1)		1.00	14.43	1.00	0.95
20.Quadratic (q=1)	0.1	100	(0.18,0.23,0.16)	(0,0,0)	1.04	1.00	1.04	0.85
		200	(0.18,0.24,0.19)	(0,0,0)	1.01	1.00	1.02	0.93
		500	(0.19,0.30,0.22)	(0,0,1)	1.01	1.00	1.00	0.98
	0.4	100	(0.17,0.19,0.16)	(0,0,0)	1.04	1.00	1.05	0.85
		200	(0.14,0.17,0.18)	(0,0,0)	1.01	1.00	1.04	0.93
		500	(0.12,0.14,0.25)	(1,0,1)	1.01	1.00	1.01	0.98
21.Quadratic (q=2)	0.1	100	(0.09,0.72,0.10,0.11,0.72,0.10,0.14,0.13,0.18)	(0,1,0,0,1,0,0,0,0)	1.99	4.25	1.20	0.66
		200	(0.08,0.92,0.10,0.10,0.92,0.10,0.16,0.12,0.21)	(0,1,0,0,1,0,0,0,0)	1.31	4.64	1.11	0.83
		500	(0.06,1,0.10,0.06,1,0.09,0.10,0.09,0.21)	(0,1,0,0,1,0,0,1,0)	1.01	4.34	1.06	0.96
	0.4	100	(0.08,0.29,0.07,0.09,0.29,0.07,0.10,0.09,0.13)	(0,1,0,0,1,0,0,0,0)	1.15	1.14	1.20	0.61
		200	(0.09,0.44,0.09,0.09,0.44,0.08,0.11,0.10,0.15)	(0,1,0,0,1,0,0,0,0)	1.10	1.15	1.11	0.80
		500	(0.07,0.66,0.10,0.06,0.66,0.06,0.10,0.10,0.22)	(0,1,0,0,1,0,0,1,0)	1.05	1.13	1.06	0.95

Continued...

Population	σ	n	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}

Table 4.3: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$, true or optimal $\boldsymbol{c}$, ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true (for linear, collinear and post-stratification functions) or optimal (for quadratic functions) estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling with $r = 0.5$ based on linear regression model for linear1-quadratic1, and quadratic model for quadratic2.

The same overall behavior can be seen in Table 4.3, which displays the expected CV and true/optimal vectors of $\boldsymbol{c}$ and MSE ratios of CV, π, and full regression estimators under the design of stratified simple random sampling without replacement with $r = 0.5$.

Population	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
1.Linear1	0	(0.18,1,0.18,0.18,0.18,0.31,0.24,0.19)	(0,1,0,0,0,0,0,0)	1.04	9.20	1.04	0.85
	0.25	(0.16,1,0.17,0.11,0.16,0.21,0.19,0.16)		1.03	10.50	1.04	0.85
	0.5	(0.12,1,0.20,0.14,0.15,0.21,0.12,0.13)		1.02	16.57	1.12	0.85
	0.75	(0.11,1,0.14,0.12,0.14,0.14,0.11,0.13)		1.05	23.01	1.27	0.84
	1	(0.10,1,0.08,0.08,0.11,0.13,0.08,0.09)		1.03	64.29	1.73	0.83
2.Linear2	0	(0.18,1,0.18,1,0.18,0.32,0.24,1)	(0,1,0,1,0,0,0,1)	1.03	41.06	1.03	0.89
	0.25	(0.15,1,0.17,1,0.16,0.21,0.19,1)		1.03	45.40	1.04	0.89
	0.5	(0.12,1,0.20,1,0.16,0.21,0.12,1)		1.02	76.84	1.08	0.89
	0.75	(0.11,1,0.14,1,0.12,0.14,0.11,1)		1.04	102.63	1.21	0.89
	1	(0.10,1,0.08,1,0.10,0.13,0.08,1)		1.03	280.11	1.51	0.88
3.Linear3	0	(1,1,0.18,1,1,0.31,0.25,1)	(1,1,0,1,1,0,0,1)	1.02	99.16	1.01	0.94
	0.25	(1,1,0.17,1,1,0.21,0.19,1)		1.02	109.69	1.02	0.93
	0.5	(1,1,0.21,1,1,0.22,0.12,1)		1.01	191.57	1.04	0.93
	0.75	(1,1,0.14,1,1,0.14,0.10,1)		1.02	251.66	1.13	0.93
	1	(1,1,0.07,1,1,0.13,0.08,1)		1.02	559.95	1.29	0.92
4.Linear4	0	(1,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.00	141.33	1.00	1.00
	0.25	(1,1,1,1,1,1,1,1)		1.00	145.72	1.00	1.00
	0.5	(1,1,1,1,1,1,1,1)		1.00	237.12	1.00	1.00
	0.75	(1,1,1,1,1,1,1,1)		1.00	308.00	1.00	1.00
	1	(1,1,1,1,1,1,1,1)		1.00	607.48	1.00	1.00
5.Collinear11	0	(0.01,1,0.24,0.40,0.17,0.18,0.16,0.16)	(0,1,0,0,0,0,0,0)	0.99	9.07	1.01	0.82
	0.25	(0.47,1,0.21,0.35,0.16,0.19,0.15,0.15)		0.99	10.43	1.01	0.84
	0.5	(0.49,1,0.16,0.38,0.16,0.15,0.14,0.14)		0.96	16.06	1.05	0.85
	0.75	(0.48,1,0.14,0.18,0.13,0.18,0.14,0.10)		0.95	26.70	1.14	0.86
	1	(0.10,1,0.09,0.09,0.09,0.12,0.13,0.11)		1.02	59.88	1.64	0.84
6.Collinear12	0	(0.01,1,0.24,1,0.17,0.18,0.16,1)	(0,1,0,1,0,0,0,1)	1.00	54.32	1.02	0.86
	0.25	(0.50,1,0.21,1,0.17,0.18,0.15,1)		0.98	60.82	1.00	0.88
	0.5	(0.53,1,0.16,1,0.15,0.16,0.13,1)		0.97	98.49	1.05	0.90
	0.75	(0.52,1,0.14,1,0.13,0.18,0.13,1)		0.96	160.21	1.09	0.90
	1	(0.09,1,0.07,1,0.09,0.12,0.12,1)		1.01	302.09	1.40	0.88
7.Collinear13	0	(1,1,0.25,1,1,0.17,0.16,1)	(1,1,0,1,1,0,0,1)	1.01	148.57	1.02	0.92
	0.25	(0,1,0.16,1,1,0.41,0.24,1)		0.64	171.91	1.01	0.89
	0.5	(0.28,1,0.26,1,1,0.39,0.20,1)		1.12	261.74	1.04	0.91
	0.75	(1,1,0.09,1,1,0.10,0.11,1)		1.01	452.10	1.09	0.93
	1	(1,1,0.08,1,1,0.11,0.11,1)		1.01	748.97	1.22	0.93
8.Collinear14	0	(1,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.00	292.17	1.00	1.00
	0.25	(0,1,1,1,1,1,1,1)		0.73	340.91	1.00	0.95
	0.5	(0.35,1,1,1,1,1,1,1)		1.15	490.75	1.00	0.96
	0.75	(1,1,1,1,1,1,1,1)		1.00	830.26	1.00	1.00
	1	(1,1,1,1,1,1,1,1)		1.00	1186.92	1.00	1.00

Continued...

Population	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
9.Collinear21	0	(0.02,1,0.22,0.39,0.16,0.17,0.16,0.16)	(0,1,0,0,0,0,0,0)	1.00	9.21	1.01	0.82
	0.25	(0.49,1,0.18,0.39,0.18,0.17,0.15,0.13)		0.98	10.26	1.01	0.84
	0.5	(0.50,1,0.16,0.41,0.17,0.15,0.14,0.15)		0.95	15.90	1.05	0.86
	0.75	(0.47,1,0.14,0.27,0.18,0.16,0.15,0.11)		0.95	27.05	1.14	0.86
	1	(0.09,1,0.08,0.09,0.09,0.12,0.13,0.12)		1.01	59.57	1.62	0.84
10.Collinear22	0	(0.02,1,0.23,1,0.16,0.18,0.17,1)	(0,1,0,1,0,0,0,1)	1.00	73.11	1.01	0.86
	0.25	(0.52,1,0.18,1,0.18,0.17,0.15,1)		0.98	81.61	1.00	0.88
	0.5	(0.56,1,0.15,1,0.17,0.16,0.15,1)		0.97	132.26	1.05	0.90
	0.75	(0.52,1,0.12,1,0.16,0.17,0.15,1)		0.95	223.59	1.08	0.90
	1	(0.06,1,0.07,1,0.08,0.11,0.13,1)		1.00	417.99	1.40	0.87
11.Collinear23	0	(1,1,0.24,1,1,0.18,0.16,1)	(1,1,0,1,1,0,0,1)	1.01	223.46	1.02	0.92
	0.25	(0,1,0.28,1,1,0.56,0.35,1)		0.68	258.67	1.01	0.90
	0.5	(0.09,1,0.39,1,1,0.62,0.37,1)		1.02	394.06	1.05	0.91
	0.75	(1,1,0.09,1,1,0.06,0.14,1)		1.01	697.08	1.09	0.93
	1	(1,1,0.08,1,1,0.11,0.12,1)		1.01	1135.66	1.22	0.93
12.Collinear24	0	(1,1,1,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.00	506.67	1.00	1.00
	0.25	(0.01,1,1,1,1,1,1,1)		0.72	595.90	1.00	0.95
	0.5	(0.09,1,1,1,1,1,1,1)		1.01	863.80	1.00	0.95
	0.75	(1,1,1,1,1,1,1,1)		1.00	1478.53	1.00	1.00
	1	(1,1,1,1,1,1,1,1)		1.00	2113.81	1.00	1.00
13.Collinear31	0	(0.06,1,0.17,0.29,0.16,0.16,0.14,0.14)	(0,1,0,0,0,0,0,0)	1.01	9.04	1.01	0.82
	0.25	(0.52,1,0.12,0.39,0.22,0.15,0.15,0.12)		0.97	9.98	1.02	0.86
	0.5	(0.55,1,0.13,0.44,0.25,0.11,0.14,0.15)		0.95	15.09	1.05	0.88
	0.75	(0.52,1,0.11,0.35,0.28,0.12,0.16,0.11)		0.95	26.90	1.16	0.87
	1	(0.09,1,0.10,0.13,0.09,0.13,0.13,0.13)		1.01	57.73	1.61	0.84
14.Collinear32	0	(0.09,1,0.20,1,0.16,0.19,0.16,1)	(0,1,0,1,0,0,0,1)	1.00	100.46	1.01	0.86
	0.25	(0.50,1,0.15,1,0.22,0.17,0.15,1)		0.98	115.81	1.01	0.89
	0.5	(0.54,1,0.13,1,0.22,0.13,0.16,1)		0.97	181.41	1.05	0.91
	0.75	(0.52,1,0.12,1,0.25,0.15,0.17,1)		0.96	319.65	1.11	0.91
	1	(0.05,1,0.09,1,0.08,0.10,0.15,1)		1.01	602.37	1.41	0.87
15.Collinear33	0	(0.64,1,0.18,1,1,0.21,0.15,1)	(1,1,0,1,1,0,0,1)	1.01	344.37	1.02	0.90
	0.25	(0,1,0.13,1,1,0.14,0.14,1)		0.86	404.14	1.02	0.88
	0.5	(0,1,0.13,1,1,0.12,0.14,1)		0.81	622.08	1.05	0.88
	0.75	(0,1,0.13,1,1,0.08,0.18,1)		0.83	1092.25	1.10	0.90
	1	(0.99,1,0.1,1,1,0.1,0.14,1)		1.00	1771.08	1.21	0.93
16.Collinear34	0	(0.98,1,0.98,1,1,1,1,1)	(1,1,1,1,1,1,1,1)	1.02	846.10	1.00	1.00
	0.25	(0.18,1,1,1,1,1,1,1)		0.89	999.63	1.00	0.96
	0.5	(0.04,1,1,1,1,1,1,1)		0.82	1491.56	1.00	0.95
	0.75	(0,1,1,1,1,1,1,1)		0.82	2501.75	1.00	0.96
	1	(0.98,1,1,1,1,1,1,1)		1.00	3651.46	1.00	1.00
17.Post-ST1	0	(0.72,0.15,0.06,0.07)	(1,0,0,0)	1.01	1.00	1.03	0.80
	0.25	(0.74,0.13,0.08,0.05)		1.02	1.00	1.09	0.81
	0.5	(0.75,0.16,0.06,0.03)		1.01	1.00	1.08	0.81
	0.75	(0.80,0.13,0.06,0.01)		1.00	1.00	1.21	0.81
	1	(0.88,0.09,0.02,0)		1.01	1.00	1.51	0.81
18.Post-ST2	0	(0,0,0.83,0.17)		1.01	36.55	1.03	0.80
	0.25	(0,0,0.87,0.13)	(0,0,1,0)	1.01	43.77	1.05	0.81
	0.5	(0,0,0.91,0.09)		1.01	60.37	1.04	0.81
	0.75	(0,0,0.94,0.06)		1.00	87.59	1.14	0.81
	1	(0,0,0.98,0.02)		1.00	179.52	1.30	0.81
19.Post-ST3	0	(0,0,0,1)		1.00	146.22	1.00	0.80
	0.25	(0,0,0,1)	(0,0,0,1)	1.00	170.54	1.00	0.81
	0.5	(0,0,0,1)		1.00	244.96	1.00	0.81
	0.75	(0,0,0,1)		1.00	320.72	1.00	0.81
	1	(0,0,0,1)		1.00	571.83	1.00	0.81
20.Quadratic1 (q=1)	0	(0.20,0.17,0.15)	(0,0,0)	1.02	1.00	1.03	0.93
	0.25	(0.19,0.21,0.18)	(0,0,0)	1.02	1.00	1.03	0.93
	0.5	(0.18,0.24,0.19)	(0,0,0)	1.01	1.00	1.02	0.93
	0.75	(0.17,0.24,0.17)	(0,0,0)	1.02	1.00	1.03	0.93

Continued. . .

Population	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
	1	(0.17,0.29,0.16)	(0,0,0)	1.03	1.00	1.04	0.93
21.Quadratic2 (q=2)	0	(0.08,0.92,0.09,0.10,0.92,0.10,0.16,0.14,0.18)	(0,1,0,0,1,0,0,0,0)	1.12	2.95	1.05	0.83
	0.25	(0.10,0.94,0.09,0.10,0.94,0.10,0.19,0.15,0.23)	(0,1,0,0,1,0,0,0,1)	1.24	3.48	1.08	0.84
	0.5	(0.08,0.92,0.10,0.10,0.92,0.10,0.16,0.12,0.21)	(0,1,0,0,1,0,0,0,0)	1.31	4.64	1.11	0.83
	0.75	(0.08,0.94,0.09,0.07,0.94,0.07,0.14,0.09,0.21)	(0,1,0,0,1,0,0,0,0)	1.50	6.81	1.25	0.83
	1	(0.06,0.92,0.08,0.07,0.92,0.08,0.11,0.07,0.19)	(0,1,0,0,1,0,0,0,0)	1.93	12.81	1.63	0.82

Table 4.4: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$, true or optimal $\boldsymbol{c}$, ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true (for linear, collinear and post-stratification functions) or optimal (for quadratic functions) estimator , and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling with model variance $\sigma^2 = 0.01$ and stratum sample sizes $(30, 40, 60, 70)$ based on linear regression model for linear1-quadratic1, and quadratic model for quadratic2.

Table 4.4 shows the expected vector of $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$, true/optimal vectors of $\boldsymbol{c}$, MSE ratios, and ratios of variance of weights of CV, π, and full regression estimators for all populations under the design of STSI with stratum sample sizes $(30, 40, 60, 70)$ and $\sigma^2 = 0.01$. The decrease of r means that the relationship between inclusion probability and model error becomes weaker. Apparently, both MSE ratios of CV and full regression estimators are stable across r, while the MSE ratio of π estimator is obviously an increasing function of r except for post-stratification1 and quadratic1 functions. The ratios of variance of weights are very stable across r for all population functions.

4.4 Conclusion

In this chapter, we proposed a design-based CV variance estimator and compared it to MSE_p. We developed theoretical results by proving that the design-based properties of the CV variance estimator hold under some appropriate assumptions. By a simulation study, we showed that the MSE based on CV estimator of $\boldsymbol{c}$ usually tracks the true/optimal MSE_p quite well. And this model selection method provides an estimator generally better than π estimator and similarly efficient to full regression estimator based on all auxiliary variables, even with a subset of control variables $\boldsymbol{x}$. And this CV estimator has less various of weights than full regression estimator. Hence, we recommend the design-based CV variance estimator for selecting the set of variables $\boldsymbol{x}$ in regression estimation.

APPENDIX A. Technical Lemmas

A.1 Lemmas I

The following lemmas prove some necessary results for Chapter 1.

Lemma A.1.1. *Let assumptions A1, A2 and A4 hold, then under assumption A3 in Chapter 1, the invertibility of* $\widehat{\boldsymbol{D}}_s$ *and* $\boldsymbol{D}_U$ *for all* $\alpha \in H_\alpha$*, we can show that*

$$\sup_{\alpha \in H_\alpha} |\boldsymbol{\beta}_U| < \infty,$$

and,

$$\sup_{\alpha \in H_\alpha} \left| \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_U \right| = o_p\left(1\right).$$

Proof of Lemma A.1.1: Suppose assumptions A1, A3 and A4 in Chapter 1 hold, let $\boldsymbol{D}_U^{-1} = \left[d_{uij}\right]_{i,j=1,\ldots,p}$ exists and $\sup_{\alpha \in H_\alpha} \max_{i,j \in \{1,2,\ldots,p\}} |d_{uij}| < \infty$, $\boldsymbol{x}_i = \left(x_{i1}, x_{i2}, \ldots, x_{ip}\right)^T$, and p is a finite number. Let $\boldsymbol{\beta}_U = \left(\beta_1, \beta_2, \ldots, \beta_p\right)^T$, then for all $\alpha \in H_\alpha$ and $k \in \{1, 2, \ldots, p\}$,

$$\begin{aligned} \sup_{\alpha \in H_\alpha} \max_{k \in \{1,2,\ldots,p\}} |\beta_k| &\leq \frac{n_N}{N^2} \sum_{i=1}^{p} \sum_{j=1}^{N} \sup_{\alpha \in H_\alpha} \max_{k,i \in \{1,2,\ldots,p\}} |d_{uki}| \max_{j \in U, i \in \{1,2,\ldots,p\}} |x_{ji}| \max_{j \in U} |y_j| \\ &< \infty. \end{aligned}$$

Next, we can write $\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_U$ as follows

$$\begin{aligned} \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_U &= \frac{n_N}{N^2} \widehat{\boldsymbol{D}}_s^{-1} \boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{Y}_s - \frac{n_N}{N^2} \boldsymbol{D}_U^{-1} \boldsymbol{X}^T \boldsymbol{Y} \\ &= \left\{ \frac{n_N}{N^2} \widehat{\boldsymbol{D}}_s^{-1} \boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{Y}_s - \frac{n_N}{N^2} \boldsymbol{D}_U^{-1} \boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{Y}_s \right\} \\ &\quad + \left\{ \frac{n_N}{N^2} \boldsymbol{D}_U^{-1} \boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{Y}_s - \frac{n_N}{N^2} \boldsymbol{D}_U^{-1} \boldsymbol{X}^T \boldsymbol{Y} \right\} \\ &= B_1 + B_2. \end{aligned}$$

It can be shown that for all $\alpha \in H_\alpha$,

$$\begin{aligned}
\mathrm{E}_p\left[\left(\widehat{\boldsymbol{D}}_s - \boldsymbol{D}_U\right)\right] &= \mathrm{E}_p\left[\frac{n_N}{N^2}\left(\boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{X}_s - \boldsymbol{X}^T\boldsymbol{X}\right)\right] \\
&= \frac{n_N}{N^2}\sum_{i\in U}\boldsymbol{x}_i\boldsymbol{x}_i^T \mathrm{E}_p\left[\left(\frac{I_i}{\pi_i}-1\right)\right] \\
&= \boldsymbol{0},
\end{aligned}$$

and from assumptions A1, A2 and A4 in Chapter 1, for the (i,j)th entry of $\left(\widehat{\boldsymbol{D}}_s - \boldsymbol{D}_U\right)$, we can show that

$$\begin{aligned}
\mathrm{E}_p\left[\left(\widehat{\boldsymbol{D}}_s - \boldsymbol{D}_U\right)_{ij}^2\right] &= \frac{n_N^2}{N^4}\sum_{k\in U} x_{ki}^2 x_{kj}^2 \mathrm{E}_p\left[\left(\frac{I_k}{\pi_k}-1\right)^2\right] \\
&\quad + \frac{n_N^2}{N^4}\sum_{(k,l)\in D_{2,N}} x_{ki}x_{kj}x_{li}x_{lj}\mathrm{E}_p\left[\left(\frac{I_k}{\pi_k}-1\right)\left(\frac{I_l}{\pi_l}-1\right)\right] \\
&= O\left(\frac{n_N^2}{N^3}\right) + O\left(\frac{n_N}{N^2}\right) \\
&= o(1).
\end{aligned}$$

Which implies that $\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{D}}_s - \boldsymbol{D}_U\right| = o_p(1)$, and $\sup_{\alpha\in H_\alpha}\left|\widehat{\boldsymbol{D}}_s^{-1} - \boldsymbol{D}_U^{-1}\right| = o_p(1)$. Similarly, we can show that,

$$\begin{aligned}
\mathrm{E}_p\left[\frac{n_N}{N^2}\left(\boldsymbol{X}_s^T\boldsymbol{W}_s\boldsymbol{Y}_s - \boldsymbol{X}^T\boldsymbol{Y}\right)\right] &= \frac{n_N}{N^2}\sum_{i\in U}\boldsymbol{x}_i y_i \mathrm{E}_p\left[\left(\frac{I_i}{\pi_i}-1\right)\right] \\
&= \boldsymbol{0},
\end{aligned}$$

and the kth entry has the following property,

$$\begin{aligned}
\mathrm{E}_p\left[\frac{n_N^2}{N^4}\left(\boldsymbol{X}_s^T\boldsymbol{W}_s\boldsymbol{Y}_s - \boldsymbol{X}^T\boldsymbol{Y}\right)_k^2\right] &= \frac{n_N^2}{N^4}\sum_{i\in U} x_{ik}^2 y_i^2 \mathrm{E}_p\left[\left(\frac{I_i}{\pi_i}-1\right)^2\right] \\
&\quad + \frac{n_N^2}{N^4}\sum_{(i,j)\in D_{2,N}} x_{ik}x_{jk}y_i y_j \mathrm{E}_p\left[\left(\frac{I_i}{\pi_i}-1\right)\left(\frac{I_j}{\pi_j}-1\right)\right] \\
&= O\left(\frac{n_N^2}{N^3}\right) + O\left(\frac{n_N}{N^2}\right) \\
&= o(1).
\end{aligned}$$

Which implies that $\frac{n_N}{N^2}\left(\boldsymbol{X}_s^T\boldsymbol{W}_s\boldsymbol{Y}_s - \boldsymbol{X}^T\boldsymbol{Y}\right) = o_p(1)$. It can be shown that

$$\begin{aligned}
\left|\frac{n_N}{N^2}\boldsymbol{X}_s^T\boldsymbol{W}_s\boldsymbol{Y}_s\right| &\le \frac{n_N}{N^2}\sum_{i\in U}\max_{i\in U}|\boldsymbol{x}_i|\max_{i\in U}|y_i|\max_{i\in U}\left|\frac{I_i}{\pi_i}\right| \\
&= O\left(\frac{n_N}{N}\right),
\end{aligned}$$

since $\max_{i \in U} \left|\frac{I_i}{\pi_i}\right| \leq \frac{1}{\lambda}$. Therefore, we can show that

$$\begin{aligned}
\sup_{\alpha \in H_\alpha} |B_1| &\leq \sup_{\alpha \in H_\alpha} \left|\widehat{\boldsymbol{D}}_s^{-1} - \boldsymbol{D}_U^{-1}\right| \left|\frac{n_N}{N^2} \boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{Y}_s\right| \\
&= o_p(1) O\left(\frac{n_N}{N}\right) \\
&= o_p(1),
\end{aligned}$$

and from the assumption of invertibility of $\boldsymbol{D}_U$,

$$\begin{aligned}
\sup_{\alpha \in H_\alpha} |B_2| &\leq \sup_{\alpha \in H_\alpha} \left|\boldsymbol{D}_U^{-1}\right| \left|\frac{n_N}{N^2} \left(\boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{Y}_s - \boldsymbol{X}^T \boldsymbol{Y}\right)\right| \\
&= O(1) o_p(1) \\
&= o_p(1).
\end{aligned}$$

Then it follows that

$$\begin{aligned}
\sup_{\alpha \in H_\alpha} \left|\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_U\right| &\leq \sup_{\alpha \in H_\alpha} |B_1| + \sup_{\alpha \in H_\alpha} |B_2| \\
&= o_p(1) + o_p(1) \\
&= o_p(1).
\end{aligned}$$

So, the results follow.

□

Lemma A.1.2. *Under assumptions A1-A5 in Chapter 1, it can be shown that*

$$\sup_{\alpha \in H_\alpha} \max_{k \in \{1,\ldots,p\}} \mathrm{E}_p\left[b_k^4\right] = O\left(\frac{1}{n_N^2}\right),$$

where b_k represents the kth element of $\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_U$. *Proof of Lemma* A.1.2: Suppose assumptions A3 and A4 in Chapter 1 hold. Let $\widehat{\boldsymbol{d}}_s = \frac{n_N}{N^2} \boldsymbol{X}_s^T \boldsymbol{W}_s \boldsymbol{Y}_s$, and $\boldsymbol{d}_U = \frac{n_N}{N^2} \boldsymbol{X}^T \boldsymbol{Y}$. The assumptions of Theorem 5.4.3 of (Fuller 1996) with $\alpha = 1$, $s = 4$ are satisfied as follows:

(i) $\mathrm{E}_p\left[\left(\widehat{\boldsymbol{D}}_s - \boldsymbol{D}_U\right)^4\right] = O\left(\frac{1}{n_N^2}\right)$, and $\mathrm{E}_p\left[\left(\widehat{\boldsymbol{d}}_s - \boldsymbol{d}_U\right)^4\right] = O\left(\frac{1}{n_N^2}\right)$ for all $j \in \{1, \ldots, p\}$.

(ii) $\left(\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_U\right)^4 = O(1)$.

(iii) $\widehat{\boldsymbol{\beta}}^{(i_1,\ldots,i_4)}$ is continuous in $\left(\widehat{\boldsymbol{D}}_s, \widehat{\boldsymbol{d}}_s\right)$ over a closed and bounded sphere S for all $n > N_0$, where

$$\widehat{\boldsymbol{\beta}}^{(i_1,\ldots,i_r)}(\boldsymbol{x}_0) = \frac{\partial^r}{\partial x_{i_1} \cdots \partial x_{i_r}} \widehat{\boldsymbol{\beta}}\Big|_{\left(\widehat{\boldsymbol{D}}_s, \widehat{\boldsymbol{d}}_s\right) = \boldsymbol{x}_0}.$$

(iv) $(\boldsymbol{D}_U, \boldsymbol{d}_U)$ is an interior point of S.

(v) There is a finite number K such that, for $n > N_0$, $\left|\widehat{\boldsymbol{\beta}}^{(i_1,\ldots,i_4)}\right| \leq K$, for all $\left(\widehat{\boldsymbol{D}}_s, \widehat{\boldsymbol{d}}_s\right) \in S$, $\left|\boldsymbol{\beta}_U^{(i_1,\ldots,i_r)}\right| \leq K$, for $r = 1, 2, 3$.

Proof: (i) The typical elements of $\boldsymbol{D}_U$ and $\boldsymbol{d}_U$ are denoted respectively by

$$d_{jj'} = \frac{n_N}{N^2}\left(\sum_{k\in U} x_{kj}x_{kj'} + \alpha I_{\{j=j', j>1+q\}}\right) \qquad d_{j0} = \frac{n_N}{N^2}\sum_{k\in U} x_{kj}y_k.$$

The typical elements of $\widehat{\boldsymbol{D}}_s$ and $\widehat{\boldsymbol{d}}_s$ are given by

$$\widehat{d}_{jj',\pi} = \frac{n_N}{N^2}\left(\sum_{k\in U} x_{kj}x_{kj'}\frac{I_k}{\pi_k} + \alpha I_{\{j=j', j>1+q\}}\right) \qquad \widehat{d}_{j0,\pi} = \frac{n_N}{N^2}\sum_{k\in U} x_{kj}y_k\frac{I_k}{\pi_k}.$$

Let assumptions A1, A2, A4 and A5 in Chapter 1 hold. Then

$$\begin{aligned}
\mathrm{E}_p\left[\left(\widehat{d}_{jj',\pi} - d_{jj'}\right)^4\right] &= \frac{n_N^4}{N^8}\sum_{i,k,l,m\in U} x_{ij}x_{kj}x_{lj}x_{mj}x_{ij'}x_{kj'}x_{lj'}x_{mj'} \\
&\quad \times \mathrm{E}_p\left[\left(\frac{I_i}{\pi_i}-1\right)\left(\frac{I_k}{\pi_k}-1\right)\left(\frac{I_l}{\pi_l}-1\right)\left(\frac{I_m}{\pi_m}-1\right)\right] \\
&\leq \frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{k\in U}\left\{\max_{k\in U, j\in\{1,\ldots,p\}}|x_{kj}|\right\}^8 \max_{k\in U}\left|\mathrm{E}_p\left[(I_k-\pi_k)^4\right]\right| \\
&\quad + \frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{(i,k)\in D_{2,N}}\left\{\max_{k\in U, j\in\{1,\ldots,p\}}|x_{kj}|\right\}^8 \\
&\quad \times \max_{(i,k)\in D_{2,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^2(I_k-\pi_k)^2\right]\right| \\
&\quad + \frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{(i,k)\in D_{2,N}}\left\{\max_{k\in U, j\in\{1,\ldots,p\}}|x_{kj}|\right\}^8 \\
&\quad \times \max_{(i,k)\in D_{2,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^3(I_k-\pi_k)\right]\right| \\
&\quad + \frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{(i,k,l)\in D_{3,N}}\left\{\max_{k\in U, j\in\{1,\ldots,p\}}|x_{kj}|\right\}^8 \\
&\quad \times \max_{(i,k,l)\in D_{3,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^2(I_k-\pi_k)(I_l-\pi_l)\right]\right| \\
&\quad + \frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{(i,k,l,m)\in D_{4,N}}\left\{\max_{k\in U, j\in\{1,\ldots,p\}}|x_{kj}|\right\}^8 \\
&\quad \times \max_{(i,k,l,m)\in D_{4,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)(I_k-\pi_k)(I_l-\pi_l)(I_m-\pi_m)\right]\right| \\
&= O\left(\frac{n_N^4}{N^7}\right) + O\left(\frac{n_N^4}{N^6}\right) + O\left(\frac{n_N^4}{N^6}\right) + O\left(\frac{n_N^3}{N^5}\right) + O\left(\frac{n_N^2}{N^4}\right) \\
&= O\left(\frac{1}{n_N^2}\right).
\end{aligned}$$

Similarly, it can be shown that

$$
\begin{aligned}
\mathrm{E}_p\left[\left(\widehat{d}_{j0,\pi}-d_{j0}\right)^4\right] &= \frac{n_N^4}{N^8}\sum_{i,k,l,m\in U} x_{ij}x_{kj}x_{lj}x_{mj}y_iy_ky_ly_m \\
&\quad \times \mathrm{E}_p\left[\left(\frac{I_i}{\pi_i}-1\right)\left(\frac{I_k}{\pi_k}-1\right)\left(\frac{I_l}{\pi_l}-1\right)\left(\frac{I_m}{\pi_m}-1\right)\right] \\
&\leq \frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{k\in U}\left\{\max_{k\in U,j\in\{1,\ldots,p\}}|x_{kj}|\right\}^4\left\{\max_{k\in U}|y_k|\right\}^4\max_{k\in U}\left|\mathrm{E}_p\left[(I_k-\pi_k)^4\right]\right| \\
&\quad +\frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{(i,k)\in D_{2,N}}\left\{\max_{k\in U,j\in\{1,\ldots,p\}}|x_{kj}|\right\}^4\left\{\max_{k\in U}|y_k|\right\}^4 \\
&\quad \times\max_{(i,k)\in D_{2,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^2(I_k-\pi_k)^2\right]\right| \\
&\quad +\frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{(i,k)\in D_{2,N}}\left\{\max_{k\in U,j\in\{1,\ldots,p\}}|x_{kj}|\right\}^4\left\{\max_{k\in U}|y_k|\right\}^4 \\
&\quad \times\max_{(i,k)\in D_{2,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^3(I_k-\pi_k)\right]\right| \\
&\quad +\frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{(i,k,l)\in D_{3,N}}\left\{\max_{k\in U,j\in\{1,\ldots,p\}}|x_{kj}|\right\}^4\left\{\max_{k\in U}|y_k|\right\}^4 \\
&\quad \times\max_{(i,k,l)\in D_{3,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)^2(I_k-\pi_k)(I_l-\pi_l)\right]\right| \\
&\quad +\frac{n_N^4}{N^8}\frac{1}{\lambda^4}\sum_{(i,k,l,m)\in D_{4,N}}\left\{\max_{k\in U,j\in\{1,\ldots,p\}}|x_{kj}|\right\}^4\left\{\max_{k\in U}|y_k|\right\}^4 \\
&\quad \times\max_{(i,k,l,m)\in D_{4,N}}\left|\mathrm{E}_p\left[(I_i-\pi_i)(I_k-\pi_k)(I_l-\pi_l)(I_m-\pi_m)\right]\right| \\
&= O\left(\frac{n_N^4}{N^7}\right)+O\left(\frac{n_N^4}{N^6}\right)+O\left(\frac{n_N^4}{N^6}\right)+O\left(\frac{n_N^3}{N^5}\right)+O\left(\frac{n_N^2}{N^4}\right) \\
&= O\left(\frac{1}{n_N^2}\right).
\end{aligned}
$$

(ii) Assume that $\widehat{\boldsymbol{D}}_s^{-1}=[d_{sij}]_{i,j=1,\cdots,p}$ exists, and $\sup_{\alpha\in H_\alpha}\max_{k,i\in\{1,\ldots,p\}}|d_{ski}-d_{uki}|=O(1)$, then by assumptions A1-A5 in Chapter 1, we can show that

$$\begin{aligned}
\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right)_k &= \frac{n_N}{N^2}\sum_{i=1}^{p}\sum_{j=1}^{N} d_{ski}x_{ji}y_j\frac{I_j}{\pi_j} - \frac{n_N}{N^2}\sum_{i=1}^{p}\sum_{j=1}^{N} d_{uki}x_{ji}y_j \\
&= \frac{n_N}{N^2}\sum_{i=1}^{p}\sum_{j=1}^{N}\left(d_{ski}-d_{uki}\right)x_{ji}y_j\frac{I_j}{\pi_j} + \frac{n_N}{N^2}\sum_{i=1}^{p}\sum_{j=1}^{N} d_{uki}x_{ji}y_j\left(\frac{I_j}{\pi_j}-1\right) \\
&\le \frac{n_N}{N^2}\sum_{i=1}^{p}\sum_{j=1}^{N}\sup_{\alpha\in H_\alpha}\max_{k,i\in\{1,\ldots,p\}}\left|d_{ski}-d_{uki}\right|\max_{j\in U,i\in\{1,\ldots,p\}}\left|x_{ji}\right|\max_{j\in U}\left|y_j\right|\frac{1}{\lambda} \\
&\quad +\frac{n_N}{N^2}\sum_{i=1}^{p}\sum_{j=1}^{N}\sup_{\alpha\in H_\alpha}\max_{k,i\in\{1,\ldots,p\}}\left|d_{uki}\right|\max_{j\in U,i\in\{1,\ldots,p\}}\left|x_{ji}\right|\max_{j\in U}\left|y_j\right|\left(\frac{1}{\lambda}+1\right) \\
&= O\left(1\right).
\end{aligned}$$

Then the result follows.

The other assumptions are all met. Since the function $\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right)^4$ and its first three derivatives with respect to the elements of $(\boldsymbol{D}_U, \boldsymbol{d}_U)$ evaluate to zero, we conclude that $\mathrm{E}_p\left[\left(\widehat{\boldsymbol{\beta}}-\boldsymbol{\beta}_U\right)^4\right] = O\left(\frac{1}{n_N^2}\right)$. The conditions are satisfied for all $\alpha \in H_\alpha$. Therefore the result follows.

□

Lemma A.1.3. *Under assumptions A1, A3 and A4, we can show that*

$$\sup_{\alpha\in H_\alpha}\max_{i\in U}\left|y_i - m_i\right| < \infty.$$

Proof of Lemma A.1.3: Suppose assumptions A1, A3 and A4 holds, then by Lemma A.1.1

$$\begin{aligned}
\sup_{\alpha\in H_\alpha}\max_{i\in U}\left|y_i-m_i\right| &\le \max_{i\in U}\left|y_i\right| + \max_{i\in U}\left|\boldsymbol{x}_i\right|^T\sup_{\alpha\in H_\alpha}\left|\boldsymbol{\beta}_U\right| \\
&< \infty.
\end{aligned}$$

□

Lemma A.1.4. *Under assumptions A1-A4, we can show that* $\sup_{\alpha\in H_\alpha}\max_{i\in U} N\left|w_{sii}\right| < C_w$.

Proof of Lemma A.1.4: Assume assumptions A1-A4 hold, $\widehat{\boldsymbol{D}}_s^{-1} = \left[d_{sij}\right]_{i,j=1,\ldots,p}$ exists and $\sup_{\alpha\in H_\alpha}\max_{i,j\in\{1,2,\ldots,p\}}\left|d_{sij}\right| < \infty$, $\boldsymbol{x}_i = \left(x_{i1}, x_{i2}, \ldots, x_{ip}\right)^T$, and p is a finite number. Then for all $\alpha \in H_\alpha$ and $i \in U$,

$$\begin{aligned}
N\left|w_{sii}\right| &= \frac{n_N}{N}\left|\boldsymbol{x}_i^T\widehat{\boldsymbol{D}}_s^{-1}\boldsymbol{x}_i/\pi_i\right| \\
&= \frac{n_N}{N}\frac{1}{\pi_i}\left|\sum_{j=1}^{p}\sum_{k=1}^{p} x_{ij}d_{sjk}x_{ik}\right| \\
&< C,
\end{aligned}$$

where $C < \infty$. Then there exists $C_w > 0$ such that

$$\sup_{\alpha \in H_\alpha} \max_{i \in U} N \left| w_{sii} \right| < C_w.$$

□

Definition A.1.1. $g_{ij}\left(\widehat{\boldsymbol{D}}_s\right) = \frac{1}{1-w_{sii}} \frac{1}{1-w_{sjj}} - 1.$

Lemma A.1.5. *Under assumptions A1-A4, it can be shown that* $\sup_{\alpha \in H_\alpha} \max_{i,j \in U} \left| g_{ij}\left(\widehat{\boldsymbol{D}}_s\right) \right| = O\left(\frac{1}{N}\right)$, $\forall\ i,\ j \in \{1, 2, \ldots, N\}$.

Proof of Lemma A.1.5: From Lemma A.1.4, we have $\sup_{\alpha \in H_\alpha} \max_{i \in U} N \left| w_{sii} \right| < C_w$, where $C_w < \infty$ is a positive constant, then

$$
\begin{aligned}
&\Rightarrow \quad \left(\frac{1}{1 + C_w/N}\right)^2 - 1 \leq g_{ij}\left(\widehat{\boldsymbol{D}}_s\right) \leq \left(\frac{1}{1 - C_w/N}\right)^2 - 1 \\
&\Rightarrow \quad \frac{-2NC_w - C_w^2}{\left(N + C_w\right)^2} \leq g_{ij}\left(\widehat{\boldsymbol{D}}_s\right) \leq \frac{2NC_w - C_w^2}{\left(N - C_w\right)^2} \\
&\Rightarrow \quad \left| g_{ij}\left(\widehat{\boldsymbol{D}}_s\right) \right| \leq \max\left(\frac{2NC_w + C_w^2}{\left(N + C_w\right)^2}, \frac{2NC_w - C_w^2}{\left(N - C_w\right)^2}\right) \\
&\Rightarrow \quad \left| g_{ij}\left(\widehat{\boldsymbol{D}}_s\right) \right| \leq \frac{2NC_w - C_w^2}{\left(N - C_w\right)^2} \\
&\Rightarrow \quad N \left| g_{ij}\left(\widehat{\boldsymbol{D}}_s\right) \right| < \frac{2C_w}{\left(1 - C_w/C_w{'}\right)^2}. \qquad \left(C_w < C_w{'} < N\right)
\end{aligned}
$$

Then, there exists a constant $C_g > 0$ such that

$$\sup_{\alpha \in H_\alpha} \max_{i,j \in U} N \left| g_{ij}\left(\widehat{\boldsymbol{D}}_s\right) \right| < C_g.$$

□

A.2 Lemmas II

The following lemmas prove some necessary results for Chapter 2 and Chapter 3.

Lemma A.2.1. *Under assumptions A1-A5 in Chapter 2 or Chapter 3, we can show that*

$$\left| \boldsymbol{\beta}_N \right| < \infty,$$

and,

$$\left| \widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_N \right| = o_p\left(1\right).$$

From assumption A5, the boundary of σ_i, the proof is almost the same as the proof of Lemma A.1.1 since $\frac{n_N}{N^2}\widehat{\boldsymbol{T}}$ and $\frac{n_N}{N^2}\boldsymbol{T}$ have similar properties to $\widehat{\boldsymbol{D}}_s$ and $\boldsymbol{D}_U$ respectively. Here we assume population moments for y_i instead of the assumption of bounded y_i.

$\square$

Lemma A.2.2. *Under assumptions A1-A6 in Chapter 2 or Chapter 3, it can be shown that*

$$\max_{k\in\{1,\ldots,J\}} \mathrm{E}_p\left[b_k^4\right] = O\left(\frac{1}{n_N^2}\right),$$

where b_k represents the kth element of $\widehat{\boldsymbol{\beta}} - \boldsymbol{\beta}_N$. The proof is similar to Lemma A.1.2.

$\square$

Lemma A.2.3. *Under assumptions A1, A3, A4 and A5 in Chapter 2, we can show that*

$$\max_{i\in U}\left|y_i - y_i^0\right| < \infty.$$

Proof of Lemma A.2.3: Suppose assumptions A1, A3, A4 and A5 in Chapter 2 holds, then by Lemma A.2.1

$$\begin{aligned}\max_{i\in U}\left|y_i - y_i^0\right| &\leq \max_{i\in U}|y_i| + \max_{i\in U}|\boldsymbol{x}_i|^T|\boldsymbol{\beta}_N| \\ &< \infty.\end{aligned}$$

$\square$

Lemma A.2.4. *Under assumptions A1, A3, A4 and A5 in Chapter 3, we can show that*

$$\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i - \gamma y_i^0\right| < \infty.$$

Proof of Lemma A.2.4: Suppose assumptions A1, A3, A4 and A5 in Chapter 3 holds, then by Lemma A.2.1

$$\begin{aligned}\sup_{\gamma\in[0,1]}\max_{i\in U}\left|y_i - \gamma y_i^0\right| &\leq \max_{i\in U}|y_i| + \max_{i\in U}|\boldsymbol{x}_i|^T|\boldsymbol{\beta}_N| \\ &< \infty.\end{aligned}$$

$\square$

Lemma A.2.5. *Under assumations A1-A5 in Chapter 2 or Chapter 3, we can show that*

$$\max_{i\in U} N\left|w_{sii}^*\right| < C_{w^*},$$

where C_{w^*} is some positive constant.

Proof of Lemma A.2.5: Assume assumations A1-A5 in Chapter 2 or Chapter 3 hold, $\left(\frac{n_N}{N^2}\widehat{\boldsymbol{T}}\right)^{-1} = [t_{sij}]_{i,j=1,\ldots,J}$ exists and $\max_{i,j\in\{1,2,\ldots,J\}} |t_{sij}| < \infty$, $\boldsymbol{x}_i = (x_{i1}, x_{i2}, \ldots, x_{iJ})^T$, and J is a finite number. Then for all $i \in U$,

$$\begin{aligned} N\left|w^*_{sii}\right| &= \frac{n_N}{N}\left|\boldsymbol{x}_i^T\left(\frac{n_N}{N^2}\widehat{\boldsymbol{T}}\right)^{-1}\boldsymbol{x}_i/\sigma_i^2\pi_i\right| \\ &= \frac{n_N}{N}\frac{1}{\sigma_i^2\pi_i}\left|\sum_{j=1}^{J}\sum_{k=1}^{J} x_{ij}t_{sjk}x_{ik}\right| \\ &< C, \end{aligned}$$

where $C < \infty$. Then there exists $C_{w^*} > 0$ such that

$$\max_{i\in U} N\left|w^*_{sii}\right| < C_{w^*}.$$

□

Definition A.2.1. $g^*_{ij}\left(\widehat{\boldsymbol{T}}\right) = \frac{1}{1-w^*_{sii}}\frac{1}{1-w^*_{sjj}} - 1.$

Lemma A.2.6. *Under assumations A1-A5 in Chapter 2 or Chapter 3, it can be shown that*

$$\max_{i,j\in U}\left|g^*_{ij}\left(\widehat{\boldsymbol{T}}\right)\right| = O\left(\frac{1}{N}\right) \quad \forall i,j \in \{1,2,\ldots,N\}.$$

The proof is almost the same as the proof of Lemma A.1.5.

□

APPENDIX B. Additional Simulation Results

B.1 More Simulation Results in Chapter 1

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
1.Linear	0	8.658	660.029	10.000	658.430
	0.25	8.515	550.058	10.000	546.854
	0.5	8.903	398.395	10.000	395.372
	0.75	9.129	273.862	10.000	272.998
	1	9.541	114.401	10.000	114.067
2.Quadratic	0	7.876	660.191	10.000	658.004
	0.25	8.110	552.195	10.000	547.551
	0.5	8.127	399.683	5.591	397.601
	0.75	7.392	276.891	10.000	275.231
	1	6.431	115.420	10.000	115.040
3.Bump	0	0.068	689.397	0.038	682.278
	0.25	0.048	604.450	0.053	599.034
	0.5	0.044	464.292	0.053	447.779
	0.75	0.035	360.985	0.076	345.637
	1	0.043	210.490	0.135	190.468
4.Jump	0	1.218	708.258	0.021	699.383
	0.25	1.033	618.955	0.385	601.013
	0.5	0.936	434.021	0.343	422.015
	0.75	0.547	333.795	0.870	320.990
	1	0.364	188.826	0.689	172.960
5.Normal CDF	0	8.575	660.865	10.000	658.670
	0.25	8.384	550.161	10.000	546.440
	0.5	8.770	396.790	10.000	395.243
	0.75	8.912	274.257	10.000	273.118
	1	9.427	114.683	10.000	114.556
6.Exponential	0	5.050	671.839	3.944	667.851
	0.25	5.206	569.589	2.477	559.386
	0.5	4.842	405.317	2.477	400.013
	0.75	3.871	286.786	4.431	283.388
	1	2.910	123.185	5.591	122.571

Continued. . .

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
7.Slow sine	0	0.401	674.343	0.433	670.622
	0.25	0.383	569.991	0.870	560.048
	0.5	0.390	412.783	0.486	409.177
	0.75	0.373	301.003	0.546	298.969
	1	0.389	143.562	0.689	142.252
8.Fast sine	0	0.004	713.167	0.004	704.336
	0.25	0.004	660.732	0.005	651.725
	0.5	0.005	538.219	0.007	519.818
	0.75	0.004	451.237	0.009	432.916
	1	0.004	373.924	0.008	338.883

Table B.1: CV smoothing parameters $\widehat{\alpha}_{\mathrm{CV}}$ and optimal smoothing parameters α_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling from population of size $N = 1000$ with model variance $\sigma^2 = 0.16$, and the stratum sample sizes $n = (30, 40, 60, 70)$.

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
1.Linear	0	8.262	95.919	10.000	94.919
	0.25	8.424	78.218	10.000	77.056
	0.5	8.509	55.090	10.000	54.552
	0.75	8.997	37.754	10.000	37.353
	1	9.564	18.933	10.000	18.909
2.Quadratic	0	2.146	97.530	2.477	96.161
	0.25	2.367	78.799	3.126	77.394
	0.5	2.392	56.826	2.205	56.353
	0.75	2.332	40.063	2.205	38.753
	1	1.874	21.896	2.477	21.604
3.Bump	0	0.005	110.832	0.002	106.495
	0.25	0.004	105.068	0.005	102.477
	0.5	0.004	88.472	0.008	82.967
	0.75	0.004	74.696	0.008	68.749
	1	0.004	68.533	0.010	62.522
4.Jump	0	0.011	130.799	0.001	123.777
	0.25	0.012	142.538	0.005	130.908
	0.5	0.012	117.315	0.007	107.400
	0.75	0.008	108.561	0.008	95.979
	1	0.006	91.295	0.017	76.334

Continued. . .

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
5.Normal CDF	0	7.439	97.065	10.000	95.853
	0.25	7.521	79.078	10.000	77.837
	0.5	7.351	55.906	10.000	55.024
	0.75	7.256	38.554	10.000	37.834
	1	7.886	19.625	10.000	19.417
6.Exponential	0	0.261	103.208	0.030	101.050
	0.25	0.285	86.558	0.242	83.425
	0.5	0.287	64.886	0.171	61.646
	0.75	0.250	46.932	0.215	44.361
	1	0.192	30.528	0.343	28.537
7.Slow sine	0	0.090	101.871	0.042	99.774
	0.25	0.089	86.807	0.095	84.961
	0.5	0.090	63.817	0.107	62.343
	0.75	0.089	46.385	0.135	45.075
	1	0.093	32.274	0.152	31.909
8.Fast sine	0	0.011	175.702	0.000	131.538
	0.25	0.011	167.182	0.001	156.116
	0.5	0.021	153.564	0.001	138.172
	0.75	0.031	146.093	0.001	126.775
	1	0.021	149.133	0.001	139.864

Table B.2: CV smoothing parameters $\widehat{\alpha}_{\mathrm{CV}}$ and optimal smoothing parameters α_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling from population of size $N = 1000$ with model variance $\sigma^2 = 0.01$, and the stratum sample sizes $n = (15, 20, 30, 35)$.

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
1.Linear	0	8.262	1534.708	10.000	1518.707
	0.25	8.424	1251.496	10.000	1232.889
	0.5	8.509	881.432	10.000	872.835
	0.75	8.997	604.063	10.000	597.646
	1	9.564	302.929	10.000	302.543
2.Quadratic	0	7.333	1532.491	10.000	1513.313
	0.25	7.625	1242.319	10.000	1222.108
	0.5	7.590	881.681	10.000	875.107
	0.75	7.315	610.898	10.000	598.253
	1	6.616	310.934	10.000	307.566
3.Bump	0	0.481	1699.317	0.017	1627.184

Continued. . .

Population	r	$\widehat{\alpha}_{\text{CV}}$	$\text{MSE}_{\widehat{\alpha}_{\text{CV}}}$	α_{opt}	$\text{MSE}_{\alpha_{\text{opt}}}$
	0.25	0.367	1472.553	0.060	1410.794
	0.5	0.194	1103.057	0.107	1028.741
	0.75	0.222	877.924	0.085	784.562
	1	0.134	593.748	0.171	542.247
4.Jump	0	3.079	1643.394	0.095	1614.419
	0.25	2.634	1399.597	0.977	1341.978
	0.5	2.303	1032.115	0.774	993.904
	0.75	1.868	776.912	0.774	732.094
	1	1.862	501.575	1.556	458.652
5.Normal CDF	0	8.134	1537.078	10.000	1521.470
	0.25	8.360	1253.047	10.000	1235.172
	0.5	8.388	887.062	10.000	873.790
	0.75	8.874	606.088	10.000	598.672
	1	9.450	304.155	10.000	303.645
6.Exponential	0	5.348	1563.635	10.000	1539.973
	0.25	5.781	1257.784	6.280	1245.424
	0.5	5.422	915.976	3.944	903.500
	0.75	4.924	643.581	4.431	620.161
	1	3.986	339.676	7.055	332.355
7.Slow sine	0	0.406	1602.087	0.171	1575.821
	0.25	0.399	1333.071	0.486	1299.721
	0.5	0.399	962.404	0.433	938.687
	0.75	0.386	677.365	0.486	660.595
	1	0.398	407.195	0.614	406.647
8.Fast sine	0	0.148	1807.160	0.002	1673.098
	0.25	0.171	1678.223	0.006	1616.895
	0.5	0.211	1599.195	0.005	1426.315
	0.75	0.205	1326.333	0.007	1159.876
	1	0.175	1217.545	0.007	1094.870

Table B.3: CV smoothing parameters $\widehat{\alpha}_{\text{CV}}$ and optimal smoothing parameters α_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling from population of size $N = 1000$ with model variance $\sigma^2 = 0.16$, and the stratum sample sizes $n = (15, 20, 30, 35)$.

Population	r	$\widehat{\alpha}_{\text{CV}}$	$\text{MSE}_{\widehat{\alpha}_{\text{CV}}}$	α_{opt}	$\text{MSE}_{\alpha_{\text{opt}}}$
1.Linear	0	8.634	10.655	10.000	10.638
	0.25	8.368	8.631	10.000	8.589

Continued...

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
	0.5	8.614	7.313	10.000	7.234
	0.75	8.935	4.671	10.000	4.637
	1	9.378	2.009	10.000	1.996
2.Quadratic	0	1.877	10.703	1.233	10.636
	0.25	1.965	8.617	1.385	8.576
	0.5	1.698	7.321	2.477	7.295
	0.75	1.102	4.831	1.963	4.790
	1	0.661	2.199	1.748	2.143
3.Bump	0	0.003	11.041	0.002	10.986
	0.25	0.003	9.295	0.004	9.259
	0.5	0.003	8.754	0.006	8.556
	0.75	0.003	6.437	0.008	6.027
	1	0.003	4.328	0.008	3.859
4.Jump	0	0.000	12.845	0.000	12.800
	0.25	0.000	11.240	0.000	11.159
	0.5	0.000	10.591	0.003	9.675
	0.75	0.000	9.103	0.002	8.354
	1	0.000	7.423	0.004	5.779
5.Normal CDF	0	5.825	10.673	2.783	10.666
	0.25	4.898	8.677	10.000	8.617
	0.5	4.691	7.383	10.000	7.283
	0.75	4.964	4.727	5.591	4.703
	1	5.228	2.067	10.000	2.047
6.Exponential	0	0.096	10.807	0.095	10.719
	0.25	0.086	8.695	0.215	8.620
	0.5	0.064	7.750	0.242	7.535
	0.75	0.057	5.249	0.095	5.208
	1	0.051	2.586	0.171	2.461
7.Slow sine	0	0.050	10.864	0.053	10.845
	0.25	0.039	8.936	0.048	8.889
	0.5	0.041	7.812	0.085	7.668
	0.75	0.045	5.270	0.095	5.141
	1	0.043	2.696	0.135	2.526
8.Fast sine	0	0.001	12.032	0.001	11.961
	0.25	0.001	10.443	0.001	10.378
	0.5	0.001	10.157	0.001	9.924
	0.75	0.001	8.386	0.002	7.875
	1	0.001	6.753	0.002	5.787

Continued...

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$

Table B.4: CV smoothing parameters $\widehat{\alpha}_{\mathrm{CV}}$ and optimal smoothing parameters α_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling from population of size $N = 1000$ with model variance $\sigma^2 = 0.01$, and the stratum sample sizes $n = (75, 100, 150, 175)$.

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
1.Linear	0	8.634	170.487	10.000	170.215
	0.25	8.368	138.101	10.000	137.425
	0.5	8.614	117.006	10.000	115.737
	0.75	8.935	74.736	10.000	74.198
	1	9.378	32.149	10.000	31.944
2.Quadratic	0	7.374	170.953	3.511	170.137
	0.25	7.571	137.772	3.944	137.282
	0.5	7.149	115.910	10.000	115.519
	0.75	6.006	74.700	10.000	74.462
	1	4.579	32.569	10.000	32.182
3.Bump	0	0.032	175.117	0.027	173.696
	0.25	0.018	143.760	0.027	143.009
	0.5	0.017	128.998	0.038	126.544
	0.75	0.013	89.112	0.060	83.940
	1	0.016	50.044	0.060	45.542
4.Jump	0	0.120	179.893	0.008	179.109
	0.25	0.154	149.862	0.042	147.325
	0.5	0.131	128.821	0.385	120.176
	0.75	0.076	91.672	0.614	85.933
	1	0.025	51.745	0.343	42.499
5.Normal CDF	0	8.456	170.625	10.000	170.284
	0.25	7.948	138.083	10.000	137.475
	0.5	8.097	117.254	10.000	115.877
	0.75	8.594	74.907	10.000	74.410
	1	9.107	32.304	10.000	32.087
6.Exponential	0	3.681	171.303	2.477	169.579
	0.25	3.871	136.754	2.205	136.082
	0.5	2.913	116.866	3.944	115.908
	0.75	1.888	76.840	5.591	76.265
	1	1.341	34.857	2.783	34.081

Continued...

Population	r	$\widehat{\alpha}_{\mathrm{CV}}$	$\mathrm{MSE}_{\widehat{\alpha}_{\mathrm{CV}}}$	α_{opt}	$\mathrm{MSE}_{\alpha_{\mathrm{opt}}}$
7.Slow sine	0	0.232	172.619	0.215	172.051
	0.25	0.202	139.795	0.242	139.861
	0.5	0.208	120.576	0.385	119.704
	0.75	0.199	79.396	0.343	78.816
	1	0.187	37.424	0.486	36.322
8.Fast sine	0	0.003	178.746	0.002	177.832
	0.25	0.003	149.755	0.004	148.515
	0.5	0.003	139.213	0.007	135.693
	0.75	0.003	102.842	0.007	97.215
	1	0.002	72.083	0.007	62.024

Table B.5: CV smoothing parameters $\widehat{\alpha}_{\mathrm{CV}}$ and optimal smoothing parameters α_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling from population of size $N = 1000$ with model variance $\sigma^2 = 0.16$, and the stratum sample sizes $n = (75, 100, 150, 175)$.

B.2 More Simulation Results in Chapter 2

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
0.1	100	0	1.020	1.045	-3.243	-2.547	1.787
		0.25	1.021	1.043	-3.275	-2.248	1.822
		0.5	1.039	1.075	-1.296	0.338	3.911
		0.75	1.074	1.098	1.022	3.857	6.546
		1	1.148	1.054	-2.554	4.495	2.860
	200	0	1.009	1.012	-2.264	-2.009	0.261
		0.25	1.018	1.049	-0.010	0.461	2.600
		0.5	1.013	1.020	-1.868	-1.150	0.695
		0.75	1.022	1.025	-1.684	-0.463	0.980
		1	1.088	1.038	1.160	4.586	3.932
	500	0	1.001	1.001	-1.123	-1.103	-0.063
		0.25	1.002	1.024	0.164	0.265	1.248
		0.5	1.001	0.990	-1.907	-1.710	-0.839
		0.75	0.998	0.987	-2.690	-2.332	-1.595
		1	1.013	1.001	-2.403	-1.377	-1.297
0.4	100	0	1.020	1.045	-3.243	-2.547	1.787
		0.25	1.021	1.043	-3.275	-2.248	1.822
		0.5	1.039	1.075	-1.296	0.338	3.911
		0.75	1.074	1.098	1.022	3.857	6.546
		1	1.148	1.054	-2.554	4.495	2.860
	200	0	1.009	1.012	-2.264	-2.009	0.261
		0.25	1.018	1.049	-0.010	0.461	2.600
		0.5	1.013	1.020	-1.868	-1.150	0.695
		0.75	1.022	1.025	-1.684	-0.463	0.980
		1	1.088	1.038	1.160	4.586	3.932
	500	0	1.001	1.001	-1.123	-1.103	-0.063
		0.25	1.002	1.024	0.164	0.265	1.248
		0.5	1.001	0.990	-1.907	-1.710	-0.839
		0.75	0.998	0.987	-2.690	-2.332	-1.595
		1	1.013	1.001	-2.403	-1.377	-1.297

Table B.6: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Linear1 function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
0.1	100	0	1.009	0.968	-8.892	-6.781	0.697
		0.25	1.000	0.950	-9.519	-6.288	0.385
		0.5	1.048	1.059	-7.144	-2.093	3.106
		0.75	1.120	1.000	-8.897	-0.139	1.265
		1	1.265	0.952	-14.933	4.027	-4.290
	200	0	1.015	1.020	-3.405	-2.464	1.561
		0.25	1.011	1.009	-3.770	-2.272	1.376
		0.5	1.005	1.002	-4.473	-2.095	0.691
		0.75	1.109	1.050	-2.870	1.664	2.412
		1	1.222	0.982	-7.206	3.477	-1.469
	500	0	1.005	0.945	-2.902	-2.794	-0.836
		0.25	1.001	0.994	-1.756	-1.445	0.434
		0.5	1.001	1.020	-1.257	-0.642	0.976
		0.75	0.985	0.950	-3.379	-2.011	-1.169
		1	0.998	0.944	-5.933	-2.364	-3.466
0.4	100	0	1.009	0.968	-8.892	-6.781	0.697
		0.25	1.000	0.950	-9.519	-6.288	0.385
		0.5	1.048	1.059	-7.144	-2.093	3.106
		0.75	1.120	1.000	-8.897	-0.139	1.265
		1	1.265	0.952	-14.933	4.027	-4.290
	200	0	1.015	1.020	-3.405	-2.464	1.561
		0.25	1.011	1.009	-3.770	-2.272	1.376
		0.5	1.005	1.002	-4.473	-2.095	0.691
		0.75	1.109	1.050	-2.870	1.664	2.412
		1	1.222	0.982	-7.206	3.477	-1.469
	500	0	1.005	0.945	-2.902	-2.794	-0.836
		0.25	1.001	0.994	-1.756	-1.445	0.434
		0.5	1.001	1.020	-1.257	-0.642	0.976
		0.75	0.985	0.950	-3.379	-2.011	-1.169
		1	0.998	0.944	-5.933	-2.364	-3.466

Table B.7: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Linear2 function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
0.1	100	0	0.916	0.866	-17.042	-11.923	5.075

Continued...

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
		0.25	0.856	0.695	-20.488	-13.029	0.046
		0.5	0.809	0.662	-22.587	-11.713	-2.661
		0.75	0.741	0.610	-29.155	-13.057	-10.603
		1	0.701	0.712	-40.485	-11.947	-24.379
	200	0	0.957	0.889	-8.792	-6.547	2.594
		0.25	0.910	0.789	-10.376	-6.796	0.416
		0.5	0.855	0.739	-11.992	-6.303	-1.445
		0.75	0.766	0.675	-17.110	-7.871	-7.026
		1	0.742	0.768	-26.296	-5.756	-17.053
	500	0	1.010	1.037	-2.631	-2.403	2.306
		0.25	0.965	0.850	-4.741	-3.981	-0.113
		0.5	0.912	0.769	-7.118	-5.519	-2.651
		0.75	0.843	0.760	-8.946	-5.738	-4.479
		1	0.754	0.845	-19.922	-11.427	-15.901
0.4	100	0	0.916	0.866	-17.042	-11.923	5.075
		0.25	0.856	0.695	-20.488	-13.029	0.046
		0.5	0.809	0.662	-22.587	-11.713	-2.661
		0.75	0.741	0.610	-29.155	-13.057	-10.603
		1	0.701	0.712	-40.485	-11.947	-24.379
	200	0	0.957	0.889	-8.792	-6.547	2.594
		0.25	0.910	0.789	-10.376	-6.796	0.416
		0.5	0.855	0.739	-11.992	-6.303	-1.445
		0.75	0.766	0.675	-17.110	-7.871	-7.026
		1	0.742	0.768	-26.296	-5.756	-17.053
	500	0	1.010	1.037	-2.631	-2.403	2.306
		0.25	0.965	0.850	-4.741	-3.981	-0.113
		0.5	0.912	0.769	-7.118	-5.519	-2.651
		0.75	0.843	0.760	-8.946	-5.738	-4.479
		1	0.754	0.845	-19.922	-11.427	-15.901

Table B.8: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Linear3 function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
0.1	100	0	0.991	0.951	-7.698	-5.311	-2.026
		0.25	1.003	0.993	-5.832	-3.440	-0.194

Continued...

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
		0.5	1.020	1.017	-4.785	-2.296	0.894
		0.75	1.054	1.099	-1.841	0.862	4.065
		1	1.015	1.011	-4.747	-2.143	1.003
	200	0	0.990	0.958	-4.267	-3.190	-1.372
		0.25	1.009	1.021	-1.917	-0.842	0.970
		0.5	0.996	0.987	-3.196	-2.131	-0.368
		0.75	1.006	1.021	-2.018	-0.910	0.871
		1	1.001	1.000	-2.492	-1.304	0.391
	500	0	0.997	0.990	-1.417	-1.198	-0.180
		0.25	0.993	0.959	-2.607	-2.405	-1.429
		0.5	0.999	1.002	-1.019	-0.817	0.164
		0.75	1.003	1.025	-0.230	-0.034	0.969
		1	0.998	0.984	-1.476	-1.270	-0.297
0.4	100	0	1.014	1.019	-4.992	-3.937	0.027
		0.25	1.020	1.046	-3.072	-1.845	1.897
		0.5	1.034	1.056	-2.317	-0.661	2.527
		0.75	1.057	1.100	1.542	3.884	6.539
		1	1.070	1.045	-2.228	1.498	2.160
	200	0	1.007	1.007	-2.633	-2.196	-0.085
		0.25	1.021	1.059	0.720	1.273	3.269
		0.5	1.009	1.012	-2.225	-1.534	0.157
		0.75	1.004	1.012	-2.608	-1.663	-0.252
		1	1.042	1.041	1.441	3.165	3.651
	500	0	0.999	0.998	-1.323	-1.258	-0.254
		0.25	0.999	0.993	-1.538	-1.433	-0.517
		0.5	1.000	1.001	-1.106	-0.957	-0.126
		0.75	0.998	0.999	-1.623	-1.412	-0.669
		1	1.003	1.001	-1.423	-1.028	-0.595

Table B.9: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Quadratic function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
0.1	100	0	1.038	1.101	-6.215	-4.475	4.572
		0.25	1.065	1.127	-5.509	-3.068	5.188
		0.5	1.038	1.024	-7.197	-3.282	2.951

Continued...

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
		0.75	1.094	1.092	-5.908	1.134	4.371
		1	1.328	0.975	-12.432	6.749	-3.414
	200	0	1.024	1.030	-3.610	-2.827	1.787
		0.25	1.053	1.073	-2.583	-1.450	2.799
		0.5	1.014	0.988	-4.042	-2.209	1.059
		0.75	1.055	1.069	-1.981	1.358	3.238
		1	1.242	0.995	-5.493	4.981	-0.807
	500	0	1.016	0.975	-2.364	-2.161	-0.075
		0.25	1.044	1.053	-0.617	-0.345	1.653
		0.5	1.016	1.065	-0.254	0.215	1.920
		0.75	1.011	0.991	-2.288	-1.363	-0.153
		1	1.011	0.962	-5.171	-1.628	-3.285
0.4	100	0	1.038	1.101	-6.215	-4.475	4.572
		0.25	1.065	1.127	-5.509	-3.068	5.188
		0.5	1.038	1.024	-7.197	-3.282	2.951
		0.75	1.094	1.092	-5.908	1.134	4.371
		1	1.328	0.975	-12.432	6.749	-3.414
	200	0	1.024	1.030	-3.610	-2.827	1.787
		0.25	1.053	1.073	-2.583	-1.450	2.799
		0.5	1.014	0.988	-4.042	-2.209	1.059
		0.75	1.055	1.069	-1.981	1.358	3.238
		1	1.242	0.995	-5.493	4.981	-0.807
	500	0	1.016	0.975	-2.364	-2.161	-0.075
		0.25	1.044	1.053	-0.617	-0.345	1.653
		0.5	1.016	1.065	-0.254	0.215	1.920
		0.75	1.011	0.991	-2.288	-1.363	-0.153
		1	1.011	0.962	-5.171	-1.628	-3.285

Table B.10: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Collinear1 function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
0.1	100	0	1.039	1.098	-6.260	-4.505	4.517
		0.25	1.064	1.127	-5.519	-3.052	5.200
		0.5	1.039	1.024	-7.217	-3.265	2.951
		0.75	1.096	1.091	-5.937	1.155	4.338

Continued...

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
		1	1.323	0.973	-12.525	6.607	-3.527
	200	0	1.024	1.029	-3.616	-2.815	1.776
		0.25	1.052	1.073	-2.583	-1.444	2.811
		0.5	1.015	0.988	-4.034	-2.193	1.077
		0.75	1.055	1.068	-2.039	1.319	3.176
		1	1.244	0.994	-5.514	4.967	-0.839
	500	0	1.017	0.974	-2.372	-2.164	-0.084
		0.25	1.043	1.052	-0.640	-0.366	1.637
		0.5	1.017	1.066	-0.242	0.230	1.939
		0.75	1.011	0.991	-2.304	-1.372	-0.168
		1	1.012	0.961	-5.189	-1.638	-3.307
0.4	100	0	1.039	1.098	-6.260	-4.505	4.517
		0.25	1.064	1.127	-5.519	-3.052	5.200
		0.5	1.039	1.024	-7.217	-3.265	2.951
		0.75	1.096	1.091	-5.937	1.155	4.338
		1	1.323	0.973	-12.525	6.607	-3.527
	200	0	1.024	1.029	-3.616	-2.815	1.776
		0.25	1.052	1.073	-2.583	-1.444	2.811
		0.5	1.015	0.988	-4.034	-2.193	1.077
		0.75	1.055	1.068	-2.039	1.319	3.176
		1	1.244	0.994	-5.514	4.967	-0.839
	500	0	1.017	0.974	-2.372	-2.164	-0.084
		0.25	1.043	1.052	-0.640	-0.366	1.637
		0.5	1.017	1.066	-0.242	0.230	1.939
		0.75	1.011	0.991	-2.304	-1.372	-0.168
		1	1.012	0.961	-5.189	-1.638	-3.307

Table B.11: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Collinear2 function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
0.1	100	0	1.039	1.095	-6.339	-4.561	4.448
		0.25	1.062	1.128	-5.540	-3.049	5.231
		0.5	1.040	1.025	-7.226	-3.238	2.981
		0.75	1.097	1.091	-5.981	1.160	4.312
		1	1.320	0.971	-12.549	6.512	-3.556

Continued...

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
	200	0	1.025	1.028	-3.635	-2.812	1.762
		0.25	1.052	1.074	-2.593	-1.449	2.831
		0.5	1.017	0.990	-4.017	-2.169	1.113
		0.75	1.055	1.067	-2.080	1.298	3.141
		1	1.245	0.993	-5.515	4.952	-0.851
	500	0	1.018	0.974	-2.388	-2.175	-0.097
		0.25	1.042	1.052	-0.671	-0.397	1.621
		0.5	1.018	1.067	-0.234	0.240	1.957
		0.75	1.012	0.990	-2.324	-1.385	-0.185
		1	1.014	0.961	-5.189	-1.638	-3.313
0.4	100	0	1.039	1.095	-6.339	-4.561	4.448
		0.25	1.062	1.128	-5.540	-3.049	5.231
		0.5	1.040	1.025	-7.226	-3.238	2.981
		0.75	1.097	1.091	-5.981	1.160	4.312
		1	1.320	0.971	-12.549	6.512	-3.556
	200	0	1.025	1.028	-3.635	-2.812	1.762
		0.25	1.052	1.074	-2.593	-1.449	2.831
		0.5	1.017	0.990	-4.017	-2.169	1.113
		0.75	1.055	1.067	-2.080	1.298	3.141
		1	1.245	0.993	-5.515	4.952	-0.851
	500	0	1.018	0.974	-2.388	-2.175	-0.097
		0.25	1.042	1.052	-0.671	-0.397	1.621
		0.5	1.018	1.067	-0.234	0.240	1.957
		0.75	1.012	0.990	-2.324	-1.385	-0.185
		1	1.014	0.961	-5.189	-1.638	-3.313

Table B.12: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Collinear3 function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
0.1	100	0	1.040	1.092	-6.458	-4.641	4.356
		0.25	1.061	1.130	-5.578	-3.056	5.279
		0.5	1.042	1.027	-7.228	-3.193	3.041
		0.75	1.098	1.090	-6.045	1.150	4.285
		1	1.317	0.970	-12.502	6.460	-3.500

Continued. . .

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
	200	0	1.025	1.027	-3.669	-2.816	1.742
		0.25	1.051	1.075	-2.615	-1.466	2.858
		0.5	1.019	0.992	-3.990	-2.130	1.170
		0.75	1.056	1.067	-2.098	1.304	3.140
		1	1.246	0.992	-5.496	4.937	-0.843
	500	0	1.020	0.973	-2.412	-2.192	-0.113
		0.25	1.041	1.051	-0.712	-0.438	1.603
		0.5	1.018	1.067	-0.233	0.244	1.970
		0.75	1.012	0.990	-2.348	-1.403	-0.204
		1	1.016	0.961	-5.166	-1.624	-3.298
0.4	100	0	1.040	1.092	-6.458	-4.641	4.356
		0.25	1.061	1.130	-5.578	-3.056	5.279
		0.5	1.042	1.027	-7.228	-3.193	3.041
		0.75	1.098	1.090	-6.045	1.150	4.285
		1	1.317	0.970	-12.502	6.460	-3.500
	200	0	1.025	1.027	-3.669	-2.816	1.742
		0.25	1.051	1.075	-2.615	-1.466	2.858
		0.5	1.019	0.992	-3.990	-2.130	1.170
		0.75	1.056	1.067	-2.098	1.304	3.140
		1	1.246	0.992	-5.496	4.937	-0.843
	500	0	1.020	0.973	-2.412	-2.192	-0.113
		0.25	1.041	1.051	-0.712	-0.438	1.603
		0.5	1.018	1.067	-0.233	0.244	1.970
		0.75	1.012	0.990	-2.348	-1.403	-0.204
		1	1.016	0.961	-5.166	-1.624	-3.298

Table B.13: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Collinear4 function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
0.1	100	0	1.041	1.084	-6.735	-4.802	4.144
		0.25	1.059	1.134	-5.702	-3.082	5.362
		0.5	1.046	1.031	-7.218	-3.061	3.194
		0.75	1.099	1.091	-6.185	1.120	4.255
		1	1.316	0.970	-12.323	6.439	-3.254
	200	0	1.028	1.024	-3.765	-2.832	1.682

Continued...

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
		0.25	1.050	1.077	-2.682	-1.507	2.910
		0.5	1.023	0.996	-3.945	-2.037	1.280
		0.75	1.058	1.069	-2.081	1.384	3.203
		1	1.243	0.991	-5.466	4.875	-0.825
	500	0	1.024	0.970	-2.469	-2.225	-0.152
		0.25	1.038	1.049	-0.808	-0.523	1.566
		0.5	1.019	1.065	-0.263	0.233	1.965
		0.75	1.011	0.989	-2.402	-1.439	-0.251
		1	1.019	0.961	-5.100	-1.592	-3.247
0.4	100	0	1.041	1.084	-6.735	-4.802	4.144
		0.25	1.059	1.134	-5.702	-3.082	5.362
		0.5	1.046	1.031	-7.218	-3.061	3.194
		0.75	1.099	1.091	-6.185	1.120	4.255
		1	1.316	0.970	-12.323	6.439	-3.254
	200	0	1.028	1.024	-3.765	-2.832	1.682
		0.25	1.050	1.077	-2.682	-1.507	2.910
		0.5	1.023	0.996	-3.945	-2.037	1.280
		0.75	1.058	1.069	-2.081	1.384	3.203
		1	1.243	0.991	-5.466	4.875	-0.825
	500	0	1.024	0.970	-2.469	-2.225	-0.152
		0.25	1.038	1.049	-0.808	-0.523	1.566
		0.5	1.019	1.065	-0.263	0.233	1.965
		0.75	1.011	0.989	-2.402	-1.439	-0.251
		1	1.019	0.961	-5.100	-1.592	-3.247

Table B.14: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Collinear5 function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	NB_{CV}
0.1	100	0	1.024	0.943	-16.257	-13.522	3.390
		0.25	1.021	0.932	-15.910	-11.133	3.826
		0.5	0.982	0.845	-18.162	-10.307	1.222
		0.75	0.925	0.733	-25.203	-13.101	-6.731
		1	0.902	0.780	-34.969	-14.220	-17.904
	200	0	1.007	0.932	-8.901	-7.363	1.204
		0.25	1.006	0.950	-8.196	-5.631	1.957

Continued. . .

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
		0.5	0.969	0.840	-10.786	-6.419	-0.857
		0.75	0.973	0.846	-12.265	-5.079	-2.178
		1	0.920	0.842	-21.805	-7.182	-12.255
	500	0	1.017	0.953	-3.919	-3.810	0.411
		0.25	0.998	0.916	-4.524	-4.005	-0.230
		0.5	0.990	0.943	-4.035	-2.879	0.314
		0.75	0.958	0.872	-6.808	-4.591	-2.466
		1	0.918	0.885	-13.146	-7.539	-8.835
0.4	100	0	1.024	0.943	-16.257	-13.522	3.390
		0.25	1.021	0.932	-15.910	-11.133	3.826
		0.5	0.982	0.845	-18.162	-10.307	1.222
		0.75	0.925	0.733	-25.203	-13.101	-6.731
		1	0.902	0.780	-34.969	-14.220	-17.904
	200	0	1.007	0.932	-8.901	-7.363	1.204
		0.25	1.006	0.950	-8.196	-5.631	1.957
		0.5	0.969	0.840	-10.786	-6.419	-0.857
		0.75	0.973	0.846	-12.265	-5.079	-2.178
		1	0.920	0.842	-21.805	-7.182	-12.255
	500	0	1.017	0.953	-3.919	-3.810	0.411
		0.25	0.998	0.916	-4.524	-4.005	-0.230
		0.5	0.990	0.943	-4.035	-2.879	0.314
		0.75	0.958	0.872	-6.808	-4.591	-2.466
		1	0.918	0.885	-13.146	-7.539	-8.835

Table B.15: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Post-stratification1 function.

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
0.1	100	0	1.005	0.862	-20.403	-17.981	3.748
		0.25	0.962	0.777	-21.954	-17.047	1.787
		0.5	0.908	0.676	-25.655	-17.701	-2.765
		0.75	0.853	0.639	-33.031	-21.136	-11.545
		1	0.790	0.723	-45.546	-26.467	-26.901
	200	0	0.999	0.912	-10.340	-8.621	2.299
		0.25	0.984	0.891	-10.391	-7.213	2.204
		0.5	0.909	0.733	-14.596	-9.345	-2.496

Continued...

σ	n	r	δ_g	δ_{CV}	NB_n	NB_g	$\mathrm{NB}_{\mathrm{CV}}$
		0.75	0.870	0.724	-18.019	-9.737	-6.016
		1	0.802	0.785	-29.153	-13.042	-18.127
	500	0	1.019	0.928	-4.698	-4.589	0.709
		0.25	0.995	0.873	-5.541	-4.938	-0.200
		0.5	0.958	0.836	-6.461	-5.024	-1.126
		0.75	0.909	0.795	-9.713	-6.921	-4.414
		1	0.854	0.848	-17.363	-10.628	-12.196
0.4	100	0	1.005	0.862	-20.403	-17.981	3.748
		0.25	0.962	0.777	-21.954	-17.047	1.787
		0.5	0.908	0.676	-25.655	-17.701	-2.765
		0.75	0.853	0.639	-33.031	-21.136	-11.545
		1	0.790	0.723	-45.546	-26.467	-26.901
	200	0	0.999	0.912	-10.340	-8.621	2.299
		0.25	0.984	0.891	-10.391	-7.213	2.204
		0.5	0.909	0.733	-14.596	-9.345	-2.496
		0.75	0.870	0.724	-18.019	-9.737	-6.016
		1	0.802	0.785	-29.153	-13.042	-18.127
	500	0	1.019	0.928	-4.698	-4.589	0.709
		0.25	0.995	0.873	-5.541	-4.938	-0.200
		0.5	0.958	0.836	-6.461	-5.024	-1.126
		0.75	0.909	0.795	-9.713	-6.921	-4.414
		1	0.854	0.848	-17.363	-10.628	-12.196

Table B.16: Ratios of RMSEs of "g-corrected" and CV variance estimators to RMSE of "naive" variance estimator, and normalized biases of variance estimators to MSE_p based on 10000 replications of stratified simple random sampling from Post-stratification2 function.

B.3 More Simulation Results in Chapter 3

n	r	R^2	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
100	0	0.01	0.263	1634.939	0.128	1666.418	1648.507	1739.112
	0	0.072	0.525	217.210	0.291	224.346	228.839	226.416
	0	0.518	0.909	16.198	0.877	16.554	33.296	16.346
	0.25	0.01	0.152	1311.809	0.101	1321.708	1316.466	1458.524
	0.25	0.072	0.434	176.714	0.291	178.082	184.504	189.886
	0.25	0.518	0.899	13.511	0.884	13.577	28.621	13.709
	0.5	0.01	0.071	842.834	0.103	855.938	843.868	1048.070
	0.5	0.072	0.354	117.671	0.345	120.960	123.332	136.449
	0.5	0.518	0.879	9.590	0.904	9.621	24.569	9.851
	0.75	0.01	0.020	577.823	0.108	588.475	577.992	811.589
	0.75	0.072	0.323	84.397	0.384	85.923	89.160	105.661
	0.75	0.518	0.879	7.353	0.922	7.199	21.801	7.628
	1	0.01	0.040	229.318	0.101	231.573	229.765	505.451
	1	0.072	0.293	39.118	0.424	39.552	43.823	65.805
	1	0.518	0.848	4.312	0.917	4.218	17.683	4.751
200	0	0.01	0.263	700.365	0.134	707.051	703.217	722.564
	0	0.072	0.687	92.947	0.446	94.741	98.491	94.071
	0	0.518	1.000	6.786	0.927	6.849	14.447	6.791
	0.25	0.01	0.172	621.607	0.116	623.102	622.786	649.922
	0.25	0.072	0.606	82.953	0.468	83.476	86.970	84.614
	0.25	0.518	0.970	6.101	0.931	6.104	12.755	6.109
	0.5	0.01	0.141	370.595	0.127	372.221	371.630	407.515
	0.5	0.072	0.535	50.556	0.532	51.178	53.979	53.055
	0.5	0.518	0.949	3.816	0.948	3.798	10.253	3.830
	0.75	0.01	0.000	253.365	0.172	259.605	253.403	321.875
	0.75	0.072	0.364	36.455	0.628	37.377	38.256	41.905
	0.75	0.518	0.909	2.967	0.963	2.948	9.198	3.025
	1	0.01	0.051	94.743	0.165	97.155	94.874	157.541
	1	0.072	0.424	15.525	0.635	15.927	18.298	20.510
	1	0.518	0.909	1.423	0.960	1.395	7.866	1.481
500	0	0.01	0.323	195.230	0.174	196.014	195.735	197.506
	0	0.072	0.687	25.475	0.655	25.624	26.606	25.713
	0	0.518	0.929	1.846	0.947	1.851	3.717	1.856
	0.25	0.01	0.000	156.079	0.133	156.880	156.082	162.749
	0.25	0.072	0.535	20.670	0.655	20.830	21.331	21.188
	0.25	0.518	0.909	1.515	0.948	1.521	3.151	1.530
	0.5	0.01	0.232	108.982	0.183	109.288	109.386	113.610
	0.5	0.072	0.707	14.556	0.749	14.552	15.866	14.791
	0.5	0.518	0.980	1.066	0.971	1.062	2.977	1.068
	0.75	0.01	0.323	71.964	0.301	72.722	72.781	75.682
	0.75	0.072	0.747	9.684	0.845	9.695	11.167	9.853
	0.75	0.518	1.000	0.710	0.987	0.707	2.611	0.711
	1	0.01	0.172	26.989	0.319	27.244	27.298	34.040
	1	0.072	0.626	4.006	0.827	4.010	5.238	4.432
	1	0.518	0.949	0.316	0.982	0.307	2.233	0.320

Table B.17: CV smoothing parameters $\widehat{\gamma}_{\text{CV}}$ and optimal smoothing parameters γ_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling for linear1 population of size $N = 1000$. MSE_0 and MSE_1 are the MSEs for the π estimator and regression estimator respectively.

n	r	R^2	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
100	0	0.01	0.152	479.863	0.150	480.432	480.251	492.069
	0	0.072	0.788	63.809	0.477	64.920	67.181	64.063
	0	0.518	1.000	4.622	0.968	4.669	9.092	4.625
	0.25	0.01	0.081	334.258	0.132	335.014	334.365	351.302
	0.25	0.072	0.505	44.348	0.441	44.863	45.768	45.736
	0.25	0.518	0.919	3.275	0.939	3.279	6.516	3.302
	0.5	0.01	0.152	256.449	0.113	258.921	256.949	272.085

Continued...

n	r	R^2	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
	0.5	0.072	0.616	34.539	0.475	35.144	36.859	35.423
	0.5	0.518	0.970	2.553	0.940	2.566	6.509	2.557
	0.75	0.01	0.091	154.641	0.136	156.582	154.821	174.381
	0.75	0.072	0.566	21.516	0.561	21.740	23.478	22.703
	0.75	0.518	0.960	1.633	0.944	1.621	5.489	1.639
	1	0.01	0.182	67.180	0.165	67.272	67.943	83.555
	1	0.072	0.596	9.783	0.640	9.655	12.104	10.878
	1	0.518	0.949	0.775	0.947	0.751	4.745	0.785
200	0	0.01	0.212	192.072	0.203	194.223	192.272	194.820
	0	0.072	0.848	25.318	0.744	25.558	26.845	25.364
	0	0.518	1.000	1.830	0.992	1.834	3.784	1.831
	0.25	0.01	0.131	139.232	0.153	140.104	139.316	142.917
	0.25	0.072	0.667	18.370	0.648	18.623	19.344	18.606
	0.25	0.518	0.949	1.338	0.964	1.346	3.083	1.343
	0.5	0.01	0.182	107.618	0.152	108.122	107.825	111.557
	0.5	0.072	0.677	14.277	0.690	14.368	15.392	14.524
	0.5	0.518	0.949	1.044	0.964	1.041	2.805	1.049
	0.75	0.01	0.182	66.626	0.195	67.269	66.836	70.959
	0.75	0.072	0.657	8.951	0.759	8.965	9.979	9.238
	0.75	0.518	0.939	0.661	0.965	0.656	2.294	0.667
	1	0.01	0.172	28.267	0.226	28.628	28.467	32.763
	1	0.072	0.687	4.030	0.803	3.994	5.206	4.265
	1	0.518	0.970	0.306	0.966	0.297	2.170	0.308
500	0	0.01	1.000	54.830	0.528	54.982	55.492	54.831
	0	0.072	1.000	7.138	0.986	7.141	7.716	7.138
	0	0.518	1.000	0.515	1.000	0.515	1.043	0.515
	0.25	0.01	0.374	39.858	0.269	40.135	39.957	40.145
	0.25	0.072	0.899	5.222	0.843	5.235	5.635	5.226
	0.25	0.518	1.000	0.377	0.979	0.377	0.878	0.377
	0.5	0.01	0.253	29.703	0.241	29.758	29.751	30.127
	0.5	0.072	0.818	3.905	0.848	3.911	4.239	3.922
	0.5	0.518	1.000	0.283	0.977	0.283	0.755	0.283
	0.75	0.01	0.152	18.187	0.352	18.283	18.206	18.742
	0.75	0.072	0.758	2.411	0.886	2.402	2.694	2.440
	0.75	0.518	0.960	0.176	0.981	0.174	0.630	0.176
	1	0.01	0.283	8.430	0.435	8.419	8.506	8.928
	1	0.072	0.859	1.152	0.896	1.137	1.538	1.162
	1	0.518	1.000	0.084	0.981	0.083	0.582	0.084

Table B.18: CV smoothing parameters $\widehat{\gamma}_{\text{CV}}$ and optimal smoothing parameters γ_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling for linear2 population of size $N = 1000$. MSE_0 and MSE_1 are the MSEs for the π estimator and regression estimator respectively.

n	r	R^2	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
100	0	0.01	0.071	308.613	0.107	311.837	308.633	312.927
	0	0.072	0.717	40.500	0.506	41.313	42.118	40.740
	0	0.518	0.949	2.935	0.961	2.950	5.812	2.941
	0.25	0.01	0.354	251.084	0.127	251.766	251.677	253.041
	0.25	0.072	0.879	32.899	0.646	33.436	35.204	32.944
	0.25	0.518	1.000	2.376	0.975	2.381	5.437	2.378
	0.5	0.01	0.030	171.780	0.132	173.547	171.787	177.227
	0.5	0.072	0.707	22.768	0.699	23.100	24.514	23.073
	0.5	0.518	0.949	1.659	0.968	1.655	4.890	1.666
	0.75	0.01	0.465	121.437	0.195	122.549	122.665	123.065
	0.75	0.072	0.919	16.002	0.804	16.130	18.761	16.022
	0.75	0.518	1.000	1.155	0.969	1.159	4.429	1.157
	1	0.01	0.303	55.075	0.385	54.924	55.575	57.718
	1	0.072	0.949	7.507	0.912	7.399	10.473	7.514
	1	0.518	1.000	0.540	0.979	0.533	4.140	0.542
200	0	0.01	0.545	136.524	0.123	137.041	136.938	136.814
	0	0.072	0.859	17.786	0.721	17.866	18.778	17.812

Continued...

n	r	R^2	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
	0	0.518	0.980	1.285	0.975	1.286	2.688	1.286
	0.25	0.01	0.424	112.250	0.210	113.087	112.524	112.752
	0.25	0.072	1.000	14.678	0.850	14.790	16.028	14.679
	0.25	0.518	1.000	1.059	0.987	1.062	2.612	1.060
	0.5	0.01	0.515	81.595	0.222	81.839	82.033	81.969
	0.5	0.072	1.000	10.671	0.881	10.718	12.017	10.672
	0.5	0.518	1.000	0.769	0.983	0.771	2.352	0.770
	0.75	0.01	0.889	55.891	0.293	56.169	56.949	55.907
	0.75	0.072	1.000	7.278	0.912	7.319	9.065	7.279
	0.75	0.518	1.000	0.525	0.983	0.527	2.211	0.525
	1	0.01	0.657	21.925	0.579	21.796	22.433	22.061
	1	0.072	1.000	2.871	0.961	2.855	4.086	2.872
	1	0.518	1.000	0.207	0.990	0.206	1.588	0.207
500	0	0.01	0.889	35.507	0.239	35.674	35.702	35.510
	0	0.072	1.000	4.623	0.882	4.649	5.000	4.623
	0	0.518	1.000	0.334	0.985	0.335	0.744	0.334
	0.25	0.01	1.000	28.739	0.406	28.892	28.980	28.739
	0.25	0.072	1.000	3.741	0.956	3.754	4.128	3.742
	0.25	0.518	1.000	0.270	0.995	0.271	0.688	0.270
	0.5	0.01	0.848	24.025	0.414	24.152	24.202	24.031
	0.5	0.072	1.000	3.128	0.946	3.137	3.469	3.129
	0.5	0.518	1.000	0.226	0.991	0.226	0.623	0.226
	0.75	0.01	0.798	15.296	0.590	15.366	15.461	15.307
	0.75	0.072	1.000	1.993	0.963	1.996	2.325	1.993
	0.75	0.518	1.000	0.144	0.993	0.144	0.522	0.144
	1	0.01	1.000	6.557	0.886	6.561	6.866	6.557
	1	0.072	1.000	0.854	0.990	0.854	1.259	0.854
	1	0.518	1.000	0.061	0.998	0.062	0.477	0.062

Table B.19: CV smoothing parameters $\widehat{\gamma}_{\text{CV}}$ and optimal smoothing parameters γ_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling for linear3 population of size $N = 1000$. MSE_0 and MSE_1 are the MSEs for the π estimator and regression estimator respectively.

n	r	R^2	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
100	0	0.01	0.000	21.395	0.077	21.415	21.395	21.642
	0	0.072	0.000	2.948	0.069	2.954	2.948	2.986
	0	0.518	0.000	0.425	0.078	0.429	0.425	0.435
	0.25	0.01	0.000	17.263	0.064	17.295	17.264	17.462
	0.25	0.072	0.000	2.390	0.064	2.399	2.390	2.420
	0.25	0.518	0.000	0.367	0.086	0.370	0.367	0.375
	0.5	0.01	0.000	12.204	0.064	12.252	12.204	12.522
	0.5	0.072	0.000	1.809	0.064	1.822	1.810	1.858
	0.5	0.518	0.000	0.363	0.096	0.368	0.363	0.373
	0.75	0.01	0.081	8.478	0.068	8.514	8.480	8.688
	0.75	0.072	0.061	1.302	0.069	1.307	1.302	1.329
	0.75	0.518	0.141	0.312	0.091	0.315	0.312	0.317
	1	0.01	0.000	3.852	0.067	3.870	3.852	4.249
	1	0.072	0.000	0.722	0.078	0.729	0.722	0.769
	1	0.518	0.212	0.278	0.120	0.283	0.278	0.283
200	0	0.01	0.717	9.540	0.069	9.557	9.564	9.544
	0	0.072	0.566	1.337	0.054	1.338	1.339	1.338
	0	0.518	0.182	0.202	0.065	0.203	0.202	0.203
	0.25	0.01	0.000	7.767	0.058	7.793	7.767	7.890
	0.25	0.072	0.000	1.102	0.048	1.105	1.102	1.118
	0.25	0.518	0.000	0.176	0.074	0.177	0.176	0.179
	0.5	0.01	0.000	5.747	0.063	5.761	5.747	5.828
	0.5	0.072	0.000	0.853	0.058	0.856	0.853	0.863
	0.5	0.518	0.162	0.161	0.092	0.162	0.162	0.163
	0.75	0.01	0.000	3.914	0.064	3.928	3.914	4.022
	0.75	0.072	0.000	0.616	0.061	0.617	0.616	0.629
	0.75	0.518	0.071	0.147	0.111	0.147	0.147	0.148
	1	0.01	0.000	1.564	0.058	1.566	1.564	1.618

Continued. . .

n	r	R^2	γ_{opt}	MSE_{opt}	$\widehat{\gamma}_{\text{CV}}$	MSE_{CV}	MSE_0	MSE_1
	1	0.072	0.000	0.293	0.072	0.293	0.293	0.302
	1	0.518	0.000	0.119	0.113	0.120	0.119	0.121
500	0	0.01	0.000	2.482	0.071	2.483	2.482	2.494
	0	0.072	0.000	0.351	0.050	0.351	0.351	0.353
	0	0.518	0.000	0.053	0.056	0.053	0.053	0.054
	0.25	0.01	0.000	2.015	0.050	2.017	2.015	2.024
	0.25	0.072	0.000	0.290	0.038	0.290	0.290	0.291
	0.25	0.518	0.000	0.049	0.100	0.049	0.049	0.049
	0.5	0.01	0.091	1.674	0.059	1.674	1.674	1.679
	0.5	0.072	0.000	0.240	0.034	0.240	0.240	0.241
	0.5	0.518	0.000	0.043	0.091	0.043	0.043	0.043
	0.75	0.01	0.000	1.075	0.064	1.074	1.075	1.084
	0.75	0.072	0.000	0.164	0.039	0.165	0.164	0.166
	0.75	0.518	0.000	0.039	0.106	0.039	0.039	0.039
	1	0.01	0.000	0.467	0.038	0.467	0.467	0.475
	1	0.072	0.000	0.085	0.055	0.085	0.085	0.086
	1	0.518	0.000	0.033	0.186	0.033	0.033	0.033

Table B.20: CV smoothing parameters $\widehat{\gamma}_{\text{CV}}$ and optimal smoothing parameters γ_{opt} with their corresponding MSEs based on 1000 replications of stratified simple random sampling for quadratic population of size $N = 1000$. MSE_0 and MSE_1 are the MSEs for the π estimator and regression estimator respectively.

B.4 More Simulation Results in Chapter 4

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.18,1,0.20,0.17,0.20,0.22,0.24,0.21)	1.04	8.65	1.06	0.73
		0.25	(0.16,1,0.16,0.17,0.16,0.19,0.16,0.17)	1.05	11.51	1.10	0.72
		0.5	(0.15,1,0.15,0.15,0.16,0.18,0.15,0.16)	1.05	15.60	1.22	0.70
		0.75	(0.14,1,0.16,0.12,0.14,0.14,0.11,0.13)	1.06	23.13	1.41	0.70
		1	(0.11,1,0.08,0.10,0.09,0.13,0.11,0.10)	1.02	50.94	2.11	0.69
	200	0	(0.18,1,0.18,0.18,0.18,0.31,0.24,0.19)	1.04	9.20	1.04	0.85
		0.25	(0.16,1,0.17,0.11,0.16,0.21,0.19,0.16)	1.03	10.50	1.04	0.85
		0.5	(0.12,1,0.20,0.14,0.15,0.21,0.12,0.13)	1.02	16.57	1.12	0.85
		0.75	(0.11,1,0.14,0.12,0.14,0.14,0.11,0.13)	1.05	23.01	1.27	0.84
		1	(0.10,1,0.08,0.08,0.11,0.13,0.08,0.09)	1.03	64.29	1.73	0.83
	500	0	(0.13,1,0.15,0.14,0.16,0.47,0.37,0.16)	1.00	9.20	1.01	0.96
		0.25	(0.11,1,0.23,0.12,0.14,0.30,0.23,0.08)	1.01	11.18	1.03	0.96
		0.5	(0.10,1,0.25,0.11,0.13,0.24,0.06,0.09)	1.01	16.06	1.05	0.96
		0.75	(0.09,1,0.19,0.10,0.10,0.17,0.04,0.18)	1.01	22.67	1.07	0.96
		1	(0.14,1,0.08,0.11,0.13,0.16,0.04,0.08)	1.01	59.34	1.27	0.96
0.4	100	0	(0.18,1,0.20,0.17,0.20,0.22,0.24,0.21)	1.04	1.39	1.06	0.73
		0.25	(0.16,1,0.16,0.17,0.16,0.19,0.16,0.17)	1.05	1.63	1.10	0.72
		0.5	(0.15,1,0.15,0.15,0.16,0.18,0.15,0.16)	1.05	1.82	1.22	0.70
		0.75	(0.14,1,0.16,0.12,0.14,0.14,0.11,0.13)	1.06	2.24	1.41	0.70
		1	(0.11,1,0.08,0.10,0.09,0.13,0.11,0.10)	1.02	3.81	2.11	0.69
	200	0	(0.18,1,0.18,0.18,0.18,0.31,0.24,0.19)	1.04	1.41	1.04	0.85
		0.25	(0.16,1,0.17,0.11,0.16,0.21,0.19,0.16)	1.03	1.55	1.04	0.85
		0.5	(0.12,1,0.20,0.14,0.15,0.21,0.12,0.13)	1.02	1.83	1.12	0.85
		0.75	(0.11,1,0.14,0.12,0.14,0.14,0.11,0.13)	1.05	2.25	1.27	0.84
		1	(0.10,1,0.08,0.08,0.11,0.13,0.08,0.09)	1.03	4.80	1.73	0.83
	500	0	(0.13,1,0.15,0.14,0.16,0.47,0.37,0.16)	1.00	1.45	1.01	0.96
		0.25	(0.11,1,0.23,0.12,0.14,0.30,0.23,0.08)	1.01	1.58	1.03	0.96
		0.5	(0.10,1,0.25,0.11,0.13,0.24,0.06,0.09)	1.01	1.84	1.05	0.96
		0.75	(0.09,1,0.19,0.10,0.10,0.17,0.04,0.18)	1.01	2.20	1.07	0.96
		1	(0.14,1,0.08,0.11,0.13,0.16,0.04,0.08)	1.01	4.64	1.27	0.96

Table B.21: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from linear1 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.19,1,0.20,1,0.20,0.23,0.24,1)	1.03	39.57	1.05	0.80
		0.25	(0.17,1,0.17,1,0.17,0.19,0.18,1)	1.04	44.14	1.08	0.80
		0.5	(0.13,1,0.15,1,0.16,0.18,0.15,1)	1.04	71.48	1.17	0.78
		0.75	(0.13,1,0.16,1,0.14,0.13,0.11,1)	1.05	98.57	1.27	0.78
		1	(0.10,1,0.08,1,0.09,0.12,0.10,1)	1.02	198.81	1.71	0.77
	200	0	(0.18,1,0.18,1,0.18,0.32,0.24,1)	1.03	41.06	1.03	0.89
		0.25	(0.15,1,0.17,1,0.16,0.21,0.19,1)	1.03	45.40	1.04	0.89
		0.5	(0.12,1,0.20,1,0.16,0.21,0.12,1)	1.02	76.84	1.08	0.89
		0.75	(0.11,1,0.14,1,0.12,0.14,0.11,1)	1.04	102.63	1.21	0.89
		1	(0.10,1,0.08,1,0.10,0.13,0.08,1)	1.03	280.11	1.51	0.88
	500	0	(0.14,1,0.15,1,0.16,0.47,0.37,1)	0.99	38.92	1.00	0.97
		0.25	(0.12,1,0.23,1,0.14,0.30,0.23,1)	1.01	49.66	1.02	0.97
		0.5	(0.10,1,0.24,1,0.13,0.24,0.06,1)	1.01	69.76	1.06	0.97
		0.75	(0.09,1,0.18,1,0.10,0.17,0.04,1)	1.01	107.13	1.06	0.97
		1	(0.13,1,0.08,1,0.13,0.16,0.04,1)	1.01	279.83	1.21	0.97
0.4	100	0	(0.19,1,0.20,1,0.20,0.23,0.24,1)	1.03	3.41	1.05	0.80
		0.25	(0.17,1,0.17,1,0.17,0.19,0.18,1)	1.04	3.67	1.08	0.80
		0.5	(0.13,1,0.15,1,0.16,0.18,0.15,1)	1.04	5.28	1.17	0.78
		0.75	(0.13,1,0.16,1,0.14,0.13,0.11,1)	1.05	6.92	1.27	0.78
		1	(0.10,1,0.08,1,0.09,0.12,0.10,1)	1.02	13.01	1.71	0.77
	200	0	(0.18,1,0.18,1,0.18,0.32,0.24,1)	1.03	3.57	1.03	0.89
		0.25	(0.15,1,0.17,1,0.16,0.21,0.19,1)	1.03	3.76	1.04	0.89

Continued. . .

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
		0.5	(0.12,1,0.20,1,0.16,0.21,0.12,1)	1.02	5.47	1.08	0.89
		0.75	(0.11,1,0.14,1,0.12,0.14,0.11,1)	1.04	7.20	1.21	0.89
		1	(0.10,1,0.08,1,0.10,0.13,0.08,1)	1.03	18.31	1.51	0.88
	500	0	(0.14,1,0.15,1,0.16,0.47,0.37,1)	0.99	3.44	1.00	0.97
		0.25	(0.12,1,0.23,1,0.14,0.30,0.23,1)	1.01	3.97	1.02	0.97
		0.5	(0.10,1,0.24,1,0.13,0.24,0.06,1)	1.01	5.11	1.06	0.97
		0.75	(0.09,1,0.18,1,0.10,0.17,0.04,1)	1.01	7.47	1.06	0.97
		1	(0.13,1,0.08,1,0.13,0.16,0.04,1)	1.01	18.64	1.21	0.97

Table B.22: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from linear2 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(1,1,0.19,1,1,0.22,0.24,1)	1.01	103.00	1.03	0.88
		0.25	(1,1,0.16,1,1,0.20,0.18,1)	1.02	110.40	1.05	0.88
		0.5	(1,1,0.15,1,1,0.18,0.14,1)	1.02	164.57	1.08	0.87
		0.75	(1,1,0.15,1,1,0.14,0.11,1)	1.02	234.39	1.16	0.87
		1	(1,1,0.07,1,1,0.11,0.10,1)	1.01	435.79	1.38	0.86
	200	0	(1,1,0.18,1,1,0.31,0.25,1)	1.02	99.16	1.01	0.94
		0.25	(1,1,0.17,1,1,0.21,0.19,1)	1.02	109.69	1.02	0.93
		0.5	(1,1,0.21,1,1,0.22,0.12,1)	1.01	191.57	1.04	0.93
		0.75	(1,1,0.14,1,1,0.14,0.10,1)	1.02	251.66	1.13	0.93
		1	(1,1,0.07,1,1,0.13,0.08,1)	1.02	559.95	1.29	0.92
	500	0	(1,1,0.15,1,1,0.47,0.37,1)	0.99	102.15	0.99	0.99
		0.25	(1,1,0.22,1,1,0.30,0.24,1)	1.01	124.26	1.01	0.99
		0.5	(1,1,0.24,1,1,0.24,0.06,1)	1.01	176.65	1.04	0.98
		0.75	(1,1,0.18,1,1,0.18,0.05,1)	1.00	263.16	1.04	0.98
		1	(1,1,0.08,1,1,0.16,0.04,1)	1.00	655.22	1.13	0.98
0.4	100	0	(0.94,1,0.18,1,1,0.23,0.24,1)	1.02	7.32	1.03	0.88
		0.25	(0.96,1,0.16,1,1,0.20,0.18,1)	1.02	7.73	1.05	0.88
		0.5	(0.98,1,0.15,1,1,0.18,0.14,1)	1.02	11.02	1.08	0.87
		0.75	(1,1,0.15,1,1,0.14,0.11,1)	1.02	15.27	1.16	0.87
		1	(0.99,1,0.07,1,1,0.11,0.10,1)	1.02	27.44	1.38	0.86
	200	0	(1,1,0.18,1,1,0.31,0.25,1)	1.02	7.05	1.01	0.94
		0.25	(1,1,0.17,1,1,0.21,0.19,1)	1.02	7.69	1.02	0.93
		0.5	(1,1,0.21,1,1,0.22,0.12,1)	1.01	12.83	1.04	0.93
		0.75	(1,1,0.14,1,1,0.14,0.10,1)	1.02	16.37	1.13	0.93
		1	(1,1,0.07,1,1,0.13,0.08,1)	1.02	35.52	1.29	0.92
	500	0	(1,1,0.15,1,1,0.47,0.37,1)	0.99	7.11	0.99	0.99
		0.25	(1,1,0.22,1,1,0.30,0.24,1)	1.01	8.28	1.01	0.99
		0.5	(1,1,0.24,1,1,0.24,0.06,1)	1.01	11.87	1.04	0.98
		0.75	(1,1,0.18,1,1,0.18,0.05,1)	1.00	17.32	1.04	0.98
		1	(1,1,0.08,1,1,0.16,0.04,1)	1.00	41.56	1.13	0.98

Table B.23: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from linear3 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(1,1,1,1,1,1,1,1)	1.00	146.43	1.00	1.00
		0.25	(1,1,1,1,1,1,1,1)	1.00	164.42	1.00	1.00
		0.5	(1,1,1,1,1,1,1,1)	1.00	215.24	1.00	1.00
		0.75	(1,1,1,1,1,1,1,1)	1.00	274.26	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	429.13	1.00	1.00
	200	0	(1,1,1,1,1,1,1,1)	1.00	141.33	1.00	1.00

Continued...

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
		0.25	(1,1,1,1,1,1,1,1)	1.00	145.72	1.00	1.00
		0.5	(1,1,1,1,1,1,1,1)	1.00	237.12	1.00	1.00
		0.75	(1,1,1,1,1,1,1,1)	1.00	308.00	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	607.48	1.00	1.00
	500	0	(1,1,1,1,1,1,1,1)	1.00	151.31	1.00	1.00
		0.25	(1,1,1,1,1,1,1,1)	1.00	161.56	1.00	1.00
		0.5	(1,1,1,1,1,1,1,1)	1.00	244.51	1.00	1.00
		0.75	(1,1,1,1,1,1,1,1)	1.00	343.41	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	801.61	1.00	1.00
0.4	100	0	(0.93,1,1,1,1,1,1,1)	1.01	10.05	1.00	1.00
		0.25	(0.96,1,1,1,1,1,1,1)	1.00	11.12	1.00	1.00
		0.5	(0.98,1,1,1,1,1,1,1)	1.00	14.13	1.00	1.00
		0.75	(1,1,1,1,1,1,1,1)	1.00	17.79	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	27.22	1.00	1.00
	200	0	(1,1,1,1,1,1,1,1)	1.00	9.78	1.00	1.00
		0.25	(1,1,1,1,1,1,1,1)	1.00	10.08	1.00	1.00
		0.5	(1,1,1,1,1,1,1,1)	1.00	15.75	1.00	1.00
		0.75	(1,1,1,1,1,1,1,1)	1.00	19.72	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	38.28	1.00	1.00
	500	0	(1,1,1,1,1,1,1,1)	1.00	10.22	1.00	1.00
		0.25	(1,1,1,1,1,1,1,1)	1.00	10.67	1.00	1.00
		0.5	(1,1,1,1,1,1,1,1)	1.00	16.01	1.00	1.00
		0.75	(1,1,1,1,1,1,1,1)	1.00	22.42	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	50.67	1.00	1.00

Table B.24: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from linear4 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.02,1,0.21,0.27,0.19,0.20,0.19,0.17)	1.01	8.53	1.05	0.67
		0.25	(0.40,1,0.18,0.29,0.17,0.17,0.16,0.17)	1.02	10.87	1.08	0.69
		0.5	(0.42,1,0.15,0.26,0.15,0.14,0.15,0.15)	1.01	14.44	1.19	0.72
		0.75	(0.46,1,0.15,0.12,0.12,0.16,0.12,0.12)	0.96	21.17	1.31	0.72
		1	(0.10,1,0.10,0.09,0.08,0.11,0.11,0.12)	1.03	51.02	1.94	0.69
	200	0	(0.01,1,0.24,0.40,0.17,0.18,0.16,0.16)	0.99	9.07	1.01	0.82
		0.25	(0.47,1,0.21,0.35,0.16,0.19,0.15,0.15)	0.99	10.43	1.01	0.84
		0.5	(0.49,1,0.16,0.38,0.16,0.15,0.14,0.14)	0.96	16.06	1.05	0.85
		0.75	(0.48,1,0.14,0.18,0.13,0.18,0.14,0.10)	0.95	26.70	1.14	0.86
		1	(0.10,1,0.09,0.09,0.09,0.12,0.13,0.11)	1.02	59.88	1.64	0.84
	500	0	(0.01,1,0.42,0.68,0.14,0.12,0.15,0.14)	1.00	8.62	1.01	0.94
		0.25	(0.58,1,0.22,0.62,0.16,0.15,0.13,0.09)	0.99	11.78	1.01	0.95
		0.5	(0.62,1,0.17,0.68,0.18,0.17,0.13,0.08)	0.96	16.12	0.99	0.97
		0.75	(0.58,1,0.12,0.26,0.15,0.29,0.16,0.09)	0.99	22.65	1.06	0.97
		1	(0.11,1,0.06,0.07,0.08,0.16,0.13,0.08)	1.01	54.19	1.27	0.96
0.4	100	0	(0.02,1,0.21,0.27,0.19,0.20,0.19,0.17)	1.01	1.44	1.05	0.67
		0.25	(0.40,1,0.18,0.29,0.17,0.17,0.16,0.17)	1.02	1.50	1.08	0.69
		0.5	(0.42,1,0.15,0.26,0.15,0.14,0.15,0.15)	1.01	1.72	1.19	0.72
		0.75	(0.46,1,0.15,0.12,0.12,0.16,0.12,0.12)	0.96	2.10	1.31	0.72
		1	(0.10,1,0.10,0.09,0.08,0.11,0.11,0.12)	1.03	4.17	1.94	0.69
	200	0	(0.01,1,0.24,0.40,0.17,0.18,0.16,0.16)	0.99	1.44	1.01	0.82
		0.25	(0.47,1,0.21,0.35,0.16,0.19,0.15,0.15)	0.99	1.51	1.01	0.84
		0.5	(0.49,1,0.16,0.38,0.16,0.15,0.14,0.14)	0.96	1.81	1.05	0.85
		0.75	(0.48,1,0.14,0.18,0.13,0.18,0.14,0.10)	0.95	2.32	1.14	0.86
		1	(0.10,1,0.09,0.09,0.09,0.12,0.13,0.11)	1.02	4.72	1.64	0.84
	500	0	(0.01,1,0.42,0.68,0.14,0.12,0.15,0.14)	1.00	1.37	1.01	0.94
		0.25	(0.58,1,0.22,0.62,0.16,0.15,0.13,0.09)	0.99	1.65	1.01	0.95
		0.5	(0.62,1,0.17,0.68,0.18,0.17,0.13,0.08)	0.96	1.75	0.99	0.97
		0.75	(0.58,1,0.12,0.26,0.15,0.29,0.16,0.09)	0.99	2.13	1.06	0.97
		1	(0.11,1,0.06,0.07,0.08,0.16,0.13,0.08)	1.01	4.20	1.27	0.96

Continued. . .

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}

Table B.25: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear11 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.02,1,0.22,1,0.19,0.20,0.19,1)	1.01	45.76	1.04	0.74
		0.25	(0.42,1,0.20,1,0.17,0.19,0.15,1)	0.98	63.62	1.03	0.77
		0.5	(0.45,1,0.14,1,0.14,0.14,0.16,1)	0.98	78.97	1.13	0.79
		0.75	(0.48,1,0.16,1,0.12,0.16,0.12,1)	0.97	116.58	1.24	0.81
		1	(0.09,1,0.09,1,0.09,0.09,0.11,1)	1.03	242.99	1.60	0.77
	200	0	(0.01,1,0.24,1,0.17,0.18,0.16,1)	1.00	54.32	1.02	0.86
		0.25	(0.50,1,0.21,1,0.17,0.18,0.15,1)	0.98	60.82	1.00	0.88
		0.5	(0.53,1,0.16,1,0.15,0.16,0.13,1)	0.97	98.49	1.05	0.90
		0.75	(0.52,1,0.14,1,0.13,0.18,0.13,1)	0.96	160.21	1.09	0.90
		1	(0.09,1,0.07,1,0.09,0.12,0.12,1)	1.01	302.09	1.40	0.88
	500	0	(0,1,0.41,1,0.13,0.12,0.15,1)	1.00	50.85	1.02	0.95
		0.25	(0.61,1,0.21,1,0.15,0.15,0.13,1)	0.98	68.29	1.00	0.96
		0.5	(0.66,1,0.17,1,0.18,0.17,0.13,1)	0.97	103.22	0.99	0.98
		0.75	(0.66,1,0.10,1,0.13,0.30,0.13,1)	0.98	130.71	1.02	0.98
		1	(0.12,1,0.06,1,0.08,0.15,0.14,1)	1.01	295.18	1.18	0.97
0.4	100	0	(0.02,1,0.22,1,0.19,0.20,0.19,1)	1.01	3.93	1.04	0.74
		0.25	(0.42,1,0.20,1,0.17,0.19,0.15,1)	0.98	4.65	1.03	0.77
		0.5	(0.45,1,0.14,1,0.14,0.14,0.16,1)	0.98	5.62	1.13	0.79
		0.75	(0.48,1,0.16,1,0.12,0.16,0.12,1)	0.97	7.85	1.24	0.81
		1	(0.09,1,0.09,1,0.09,0.09,0.11,1)	1.03	15.91	1.60	0.77
	200	0	(0.01,1,0.24,1,0.17,0.18,0.16,1)	1.00	4.40	1.02	0.86
		0.25	(0.50,1,0.21,1,0.17,0.18,0.15,1)	0.98	4.77	1.00	0.88
		0.5	(0.53,1,0.16,1,0.15,0.16,0.13,1)	0.97	7.07	1.05	0.90
		0.75	(0.52,1,0.14,1,0.13,0.18,0.13,1)	0.96	10.48	1.09	0.90
		1	(0.09,1,0.07,1,0.09,0.12,0.12,1)	1.01	19.82	1.40	0.88
	500	0	(0,1,0.41,1,0.13,0.12,0.15,1)	1.00	4.13	1.02	0.95
		0.25	(0.61,1,0.21,1,0.15,0.15,0.13,1)	0.98	5.23	1.00	0.96
		0.5	(0.66,1,0.17,1,0.18,0.17,0.13,1)	0.97	7.23	0.99	0.98
		0.75	(0.66,1,0.10,1,0.13,0.30,0.13,1)	0.98	8.68	1.02	0.98
		1	(0.12,1,0.06,1,0.08,0.15,0.14,1)	1.01	19.28	1.18	0.97

Table B.26: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear12 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.97,1,0.22,1,1,0.22,0.18,1)	1.02	126.43	1.02	0.85
		0.25	(0.06,1,0.15,1,1,0.26,0.15,1)	0.76	172.54	1.02	0.79
		0.5	(0.43,1,0.17,1,1,0.23,0.16,1)	1.05	208.88	1.10	0.83
		0.75	(1,1,0.11,1,1,0.09,0.10,1)	1.03	355.39	1.21	0.86
		1	(1,1,0.09,1,1,0.08,0.09,1)	1.02	587.40	1.34	0.85
	200	0	(1,1,0.25,1,1,0.17,0.16,1)	1.01	148.57	1.02	0.92
		0.25	(0,1,0.16,1,1,0.41,0.24,1)	0.64	171.91	1.01	0.89
		0.5	(0.28,1,0.26,1,1,0.39,0.20,1)	1.12	261.74	1.04	0.91
		0.75	(1,1,0.09,1,1,0.10,0.11,1)	1.01	452.10	1.09	0.93
		1	(1,1,0.08,1,1,0.11,0.11,1)	1.01	748.97	1.22	0.93
	500	0	(1,1,0.41,1,1,0.12,0.14,1)	1.00	143.68	1.01	0.98
		0.25	(0,1,0.25,1,1,0.72,0.34,1)	0.63	189.56	1.01	0.96
		0.5	(0.10,1,0.47,1,1,0.80,0.34,1)	1.07	296.16	1.02	0.97
		0.75	(1,1,0.06,1,1,0.18,0.12,1)	1.01	403.93	1.04	0.98
		1	(1,1,0.06,1,1,0.13,0.14,1)	1.01	810.79	1.10	0.98

Continued. . .

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.4	100	0	(0.30,1,0.20,1,1,0.21,0.18,1)	1.02	9.12	1.02	0.80
		0.25	(0,1,0.17,1,1,0.15,0.13,1)	0.77	11.06	1.02	0.78
		0.5	(0,1,0.13,1,1,0.11,0.13,1)	0.69	13.13	1.10	0.79
		0.75	(0,1,0.13,1,1,0.12,0.10,1)	0.69	22.31	1.21	0.81
		1	(0.99,1,0.09,1,1,0.08,0.09,1)	1.02	37.27	1.34	0.85
	200	0	(0.44,1,0.20,1,1,0.18,0.15,1)	1.01	10.22	1.02	0.89
		0.25	(0,1,0.16,1,1,0.15,0.12,1)	0.76	11.59	1.01	0.88
		0.5	(0,1,0.13,1,1,0.12,0.10,1)	0.67	16.89	1.04	0.88
		0.75	(0,1,0.11,1,1,0.10,0.11,1)	0.71	28.10	1.09	0.89
		1	(1,1,0.08,1,1,0.11,0.11,1)	1.01	47.54	1.22	0.93
	500	0	(0.72,1,0.33,1,1,0.14,0.14,1)	1.02	9.99	1.01	0.97
		0.25	(0,1,0.13,1,1,0.13,0.10,1)	0.74	12.76	1.01	0.95
		0.5	(0,1,0.14,1,1,0.11,0.10,1)	0.66	18.80	1.02	0.96
		0.75	(0,1,0.10,1,1,0.13,0.12,1)	0.66	24.97	1.04	0.97
		1	(1,1,0.06,1,1,0.13,0.14,1)	1.01	51.43	1.10	0.98

Table B.27: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear13 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.99,1,1,1,1,1,1,1)	1.01	242.12	1.00	1.00
		0.25	(0.13,1,1,1,1,1,1,1)	0.94	324.73	1.00	0.92
		0.5	(0.60,1,1,1,1,1,1,1)	1.06	382.33	1.00	0.97
		0.75	(1,1,1,1,1,1,1,1)	1.00	566.10	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	881.90	1.00	1.00
	200	0	(1,1,1,1,1,1,1,1)	1.00	292.17	1.00	1.00
		0.25	(0,1,1,1,1,1,1,1)	0.73	340.91	1.00	0.95
		0.5	(0.35,1,1,1,1,1,1,1)	1.15	490.75	1.00	0.96
		0.75	(1,1,1,1,1,1,1,1)	1.00	830.26	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	1186.92	1.00	1.00
	500	0	(1,1,1,1,1,1,1,1)	1.00	272.77	1.00	1.00
		0.25	(0,1,1,1,1,1,1,1)	0.67	369.81	1.00	0.97
		0.5	(0.08,1,1,1,1,1,1,1)	1.11	561.78	1.00	0.98
		0.75	(1,1,1,1,1,1,1,1)	1.00	810.96	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	1425.82	1.00	1.00
0.4	100	0	(0.34,1,0.98,1,1,1,1,1)	1.02	16.36	1.00	0.94
		0.25	(0,1,1,1,1,1,1,1)	0.79	20.30	1.00	0.90
		0.5	(0,1,1,1,1,1,1,1)	0.73	23.74	1.00	0.92
		0.75	(0,1,1,1,1,1,1,1)	0.73	34.96	1.00	0.94
		1	(0.99,1,1,1,1,1,1,1)	1.00	55.56	1.00	1.00
	200	0	(0.50,1,1,1,1,1,1,1)	1.00	19.01	1.00	0.97
		0.25	(0,1,1,1,1,1,1,1)	0.77	21.85	1.00	0.95
		0.5	(0,1,1,1,1,1,1,1)	0.69	30.88	1.00	0.95
		0.75	(0,1,1,1,1,1,1,1)	0.75	51.06	1.00	0.96
		1	(1,1,1,1,1,1,1,1)	1.00	74.80	1.00	1.00
	500	0	(0.81,1,1,1,1,1,1,1)	1.01	17.95	1.00	0.99
		0.25	(0,1,1,1,1,1,1,1)	0.75	23.72	1.00	0.97
		0.5	(0,1,1,1,1,1,1,1)	0.67	34.89	1.00	0.98
		0.75	(0,1,1,1,1,1,1,1)	0.66	49.79	1.00	0.99
		1	(1,1,1,1,1,1,1,1)	1.00	89.53	1.00	1.00

Table B.28: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear14 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.04,1,0.21,0.26,0.18,0.19,0.18,0.17)	1.01	8.50	1.05	0.67
		0.25	(0.40,1,0.17,0.31,0.18,0.17,0.13,0.15)	1.01	10.83	1.09	0.69
		0.5	(0.44,1,0.13,0.29,0.16,0.12,0.15,0.15)	0.99	14.04	1.20	0.72
		0.75	(0.46,1,0.15,0.19,0.14,0.15,0.13,0.13)	0.95	20.73	1.31	0.73
		1	(0.10,1,0.11,0.11,0.09,0.11,0.12,0.14)	1.04	51.56	1.94	0.69
	200	0	(0.02,1,0.22,0.39,0.16,0.17,0.16,0.16)	1.00	9.21	1.01	0.82
		0.25	(0.49,1,0.18,0.39,0.18,0.17,0.15,0.13)	0.98	10.26	1.01	0.84
		0.5	(0.50,1,0.16,0.41,0.17,0.15,0.14,0.15)	0.95	15.90	1.05	0.86
		0.75	(0.47,1,0.14,0.27,0.18,0.16,0.15,0.11)	0.95	27.05	1.14	0.86
		1	(0.09,1,0.08,0.09,0.09,0.12,0.13,0.12)	1.01	59.57	1.62	0.84
	500	0	(0.02,1,0.37,0.65,0.13,0.12,0.15,0.12)	1.00	8.75	1.02	0.94
		0.25	(0.58,1,0.17,0.64,0.18,0.14,0.12,0.09)	0.99	11.66	1.01	0.96
		0.5	(0.62,1,0.14,0.68,0.20,0.17,0.12,0.07)	0.96	16.59	1.00	0.97
		0.75	(0.55,1,0.12,0.44,0.20,0.27,0.14,0.09)	0.98	22.76	1.06	0.97
		1	(0.11,1,0.09,0.09,0.09,0.17,0.14,0.10)	1.01	54.64	1.28	0.96
0.4	100	0	(0.04,0.99,0.21,0.27,0.18,0.19,0.19,0.17)	1.01	1.45	1.05	0.67
		0.25	(0.40,1,0.17,0.31,0.18,0.17,0.13,0.15)	1.01	1.44	1.09	0.69
		0.5	(0.44,1,0.13,0.29,0.16,0.12,0.15,0.15)	0.99	1.63	1.20	0.72
		0.75	(0.46,1,0.15,0.19,0.14,0.15,0.13,0.13)	0.95	1.97	1.31	0.73
		1	(0.10,1,0.11,0.11,0.09,0.11,0.12,0.14)	1.04	4.17	1.94	0.69
	200	0	(0.02,1,0.22,0.39,0.16,0.17,0.16,0.16)	1.00	1.45	1.01	0.82
		0.25	(0.49,1,0.18,0.39,0.18,0.17,0.15,0.13)	0.98	1.45	1.01	0.84
		0.5	(0.50,1,0.16,0.41,0.17,0.15,0.14,0.15)	0.95	1.74	1.05	0.86
		0.75	(0.47,1,0.14,0.27,0.18,0.16,0.15,0.11)	0.95	2.25	1.14	0.86
		1	(0.09,1,0.08,0.09,0.09,0.12,0.13,0.12)	1.01	4.70	1.62	0.84
	500	0	(0.02,1,0.37,0.65,0.13,0.12,0.15,0.12)	1.00	1.39	1.02	0.94
		0.25	(0.58,1,0.17,0.64,0.18,0.14,0.12,0.09)	0.99	1.60	1.01	0.96
		0.5	(0.62,1,0.14,0.68,0.20,0.17,0.12,0.07)	0.96	1.72	1.00	0.97
		0.75	(0.55,1,0.12,0.44,0.20,0.27,0.14,0.09)	0.98	2.03	1.06	0.97
		1	(0.11,1,0.09,0.09,0.09,0.17,0.14,0.10)	1.01	4.22	1.28	0.96

Table B.29: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear21 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.04,1,0.23,1,0.18,0.21,0.20,1)	1.02	61.62	1.05	0.74
		0.25	(0.42,1,0.18,1,0.18,0.18,0.14,1)	0.98	84.13	1.03	0.77
		0.5	(0.46,1,0.14,1,0.16,0.14,0.15,1)	0.98	107.45	1.14	0.79
		0.75	(0.48,1,0.14,1,0.14,0.17,0.13,1)	0.97	162.63	1.25	0.81
		1	(0.08,1,0.09,1,0.08,0.08,0.10,1)	1.04	347.89	1.60	0.76
	200	0	(0.02,1,0.23,1,0.16,0.18,0.17,1)	1.00	73.11	1.01	0.86
		0.25	(0.52,1,0.18,1,0.18,0.17,0.15,1)	0.98	81.61	1.00	0.88
		0.5	(0.56,1,0.15,1,0.17,0.16,0.15,1)	0.97	132.26	1.05	0.90
		0.75	(0.52,1,0.12,1,0.16,0.17,0.15,1)	0.95	223.59	1.08	0.90
		1	(0.06,1,0.07,1,0.08,0.11,0.13,1)	1.00	417.99	1.40	0.87
	500	0	(0.01,1,0.38,1,0.12,0.13,0.16,1)	1.00	68.07	1.02	0.95
		0.25	(0.64,1,0.16,1,0.17,0.14,0.14,1)	0.98	91.41	0.99	0.97
		0.5	(0.68,1,0.13,1,0.21,0.16,0.13,1)	0.96	140.38	0.99	0.98
		0.75	(0.69,1,0.10,1,0.19,0.23,0.14,1)	0.97	185.46	1.02	0.98
		1	(0.10,1,0.05,1,0.07,0.16,0.17,1)	1.01	412.25	1.19	0.97
0.4	100	0	(0.04,0.99,0.23,1,0.18,0.21,0.20,1)	1.02	4.93	1.05	0.74
		0.25	(0.42,1,0.18,1,0.18,0.18,0.14,1)	0.99	5.74	1.03	0.77
		0.5	(0.46,1,0.14,1,0.15,0.14,0.15,1)	0.98	7.15	1.14	0.79
		0.75	(0.48,1,0.14,1,0.14,0.17,0.13,1)	0.97	10.37	1.25	0.81
		1	(0.08,1,0.09,1,0.08,0.08,0.10,1)	1.04	22.43	1.60	0.76
	200	0	(0.02,1,0.23,1,0.16,0.18,0.17,1)	1.00	5.52	1.01	0.86
		0.25	(0.52,1,0.18,1,0.18,0.17,0.15,1)	0.98	5.89	1.00	0.88
		0.5	(0.56,1,0.15,1,0.17,0.16,0.15,1)	0.97	8.92	1.05	0.90
		0.75	(0.52,1,0.12,1,0.16,0.17,0.15,1)	0.95	14.03	1.08	0.90
		1	(0.06,1,0.07,1,0.08,0.11,0.13,1)	1.00	27.07	1.40	0.87
	500	0	(0.01,1,0.38,1,0.12,0.13,0.16,1)	1.00	5.18	1.02	0.95
		0.25	(0.64,1,0.16,1,0.17,0.14,0.14,1)	0.98	6.51	0.99	0.97

Continued...

σ	n	r	$\mathrm{E}(\widehat{\mathbf{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
		0.5	(0.68,1,0.13,1,0.21,0.16,0.13,1)	0.96	9.23	0.99	0.98
		0.75	(0.69,1,0.10,1,0.19,0.23,0.14,1)	0.97	11.69	1.02	0.98
		1	(0.10,1,0.05,1,0.07,0.16,0.17,1)	1.01	26.54	1.19	0.97

Table B.30: Expected $\widehat{\mathbf{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear22 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\mathbf{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.96,1,0.22,1,1,0.23,0.18,1)	1.03	185.84	1.02	0.85
		0.25	(0.05,1,0.23,1,1,0.36,0.21,1)	0.81	253.56	1.03	0.80
		0.5	(0.20,1,0.25,1,1,0.36,0.25,1)	0.99	318.63	1.11	0.82
		0.75	(1,1,0.10,1,1,0.09,0.12,1)	1.04	531.22	1.22	0.86
		1	(1,1,0.10,1,1,0.09,0.10,1)	1.03	920.64	1.34	0.85
	200	0	(1,1,0.24,1,1,0.18,0.16,1)	1.01	223.46	1.02	0.92
		0.25	(0,1,0.28,1,1,0.56,0.35,1)	0.68	258.67	1.01	0.90
		0.5	(0.09,1,0.39,1,1,0.62,0.37,1)	1.02	394.06	1.05	0.91
		0.75	(1,1,0.09,1,1,0.06,0.14,1)	1.01	697.08	1.09	0.93
		1	(1,1,0.08,1,1,0.11,0.12,1)	1.01	1135.66	1.22	0.93
	500	0	(1,1,0.38,1,1,0.13,0.15,1)	1.00	210.19	1.01	0.98
		0.25	(0,1,0.45,1,1,0.86,0.59,1)	0.69	283.63	1.01	0.97
		0.5	(0,1,0.68,1,1,0.96,0.68,1)	0.94	444.95	1.02	0.98
		0.75	(1,1,0.05,1,1,0.08,0.14,1)	1.00	632.23	1.05	0.98
		1	(1,1,0.06,1,1,0.15,0.16,1)	1.01	1226.14	1.10	0.98
0.4	100	0	(0.29,0.99,0.21,1,1,0.22,0.18,1)	1.03	12.84	1.02	0.81
		0.25	(0,1,0.16,1,1,0.16,0.14,1)	0.83	15.82	1.03	0.78
		0.5	(0,1,0.14,1,1,0.12,0.13,1)	0.75	19.62	1.11	0.79
		0.75	(0,1,0.12,1,1,0.11,0.10,1)	0.75	32.69	1.22	0.81
		1	(0.96,1,0.10,1,1,0.09,0.10,1)	1.02	58.03	1.34	0.85
	200	0	(0.44,1,0.18,1,1,0.20,0.14,1)	1.01	14.83	1.02	0.89
		0.25	(0,1,0.15,1,1,0.15,0.14,1)	0.81	16.71	1.01	0.88
		0.5	(0,1,0.12,1,1,0.12,0.11,1)	0.73	24.75	1.05	0.88
		0.75	(0,1,0.12,1,1,0.10,0.13,1)	0.77	42.71	1.09	0.89
		1	(1,1,0.08,1,1,0.11,0.12,1)	1.01	71.75	1.22	0.93
	500	0	(0.74,1,0.31,1,1,0.14,0.14,1)	1.02	14.07	1.01	0.97
		0.25	(0,1,0.11,1,1,0.13,0.11,1)	0.80	18.33	1.01	0.95
		0.5	(0,1,0.12,1,1,0.11,0.12,1)	0.74	27.58	1.02	0.96
		0.75	(0,1,0.12,1,1,0.12,0.18,1)	0.72	38.55	1.05	0.97
		1	(1,1,0.06,1,1,0.15,0.16,1)	1.01	77.23	1.10	0.98

Table B.31: Expected $\widehat{\mathbf{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear23 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\mathbf{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.98,1,1,1,1,1,1,1)	1.01	416.23	1.00	1.00
		0.25	(0.12,1,1,1,1,1,1,1)	0.91	556.63	1.00	0.91
		0.5	(0.29,1,1,1,1,1,1,1)	0.99	673.17	1.00	0.94
		0.75	(1,1,1,1,1,1,1,1)	1.00	992.51	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	1597.06	1.00	1.00
	200	0	(1,1,1,1,1,1,1,1)	1.00	506.67	1.00	1.00
		0.25	(0.01,1,1,1,1,1,1,1)	0.72	595.90	1.00	0.95
		0.5	(0.09,1,1,1,1,1,1,1)	1.01	863.80	1.00	0.95
		0.75	(1,1,1,1,1,1,1,1)	1.00	1478.53	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	2113.81	1.00	1.00
	500	0	(1,1,1,1,1,1,1,1)	1.00	471.05	1.00	1.00

Continued. . .

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
		0.25	(0,1,1,1,1,1,1,1)	0.69	645.10	1.00	0.97
		0.5	(0,1,1,1,1,1,1,1)	0.94	984.91	1.00	0.98
		0.75	(1,1,1,1,1,1,1,1)	1.00	1430.59	1.00	1.00
		1	(1,1,1,1,1,1,1,1)	1.00	2519.08	1.00	1.00
0.4	100	0	(0.33,0.99,0.94,1,1,1,1,1)	1.02	27.25	1.00	0.93
		0.25	(0.01,1,0.99,1,1,1,1,1)	0.83	34.28	1.00	0.90
		0.5	(0,1,1,1,1,1,1,1)	0.77	41.34	1.00	0.92
		0.75	(0,1,1,1,1,1,1,1)	0.77	60.83	1.00	0.94
		1	(0.96,1,1,1,1,1,1,1)	1.00	100.14	1.00	1.00
	200	0	(0.5,1,1,1,1,1,1,1)	1.00	32.31	1.00	0.97
		0.25	(0,1,1,1,1,1,1,1)	0.81	37.34	1.00	0.95
		0.5	(0,1,1,1,1,1,1,1)	0.75	53.56	1.00	0.95
		0.75	(0,1,1,1,1,1,1,1)	0.78	90.56	1.00	0.96
		1	(1,1,1,1,1,1,1,1)	1.00	132.73	1.00	1.00
	500	0	(0.8,1,1,1,1,1,1,1)	1.01	30.24	1.00	1.00
		0.25	(0,1,1,1,1,1,1,1)	0.80	40.50	1.00	0.97
		0.5	(0,1,1,1,1,1,1,1)	0.74	60.55	1.00	0.98
		0.75	(0,1,1,1,1,1,1,1)	0.72	87.44	1.00	0.99
		1	(1,1,1,1,1,1,1,1)	1.00	157.65	1.00	1.00

Table B.32: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear24 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.08,0.96,0.20,0.21,0.18,0.17,0.18,0.13)	1.02	8.18	1.05	0.67
		0.25	(0.40,0.99,0.14,0.29,0.20,0.13,0.13,0.14)	1.01	9.85	1.12	0.70
		0.5	(0.48,1,0.13,0.28,0.21,0.12,0.15,0.14)	1.00	13.10	1.22	0.73
		0.75	(0.49,1,0.13,0.27,0.18,0.12,0.16,0.13)	0.96	20.36	1.34	0.74
		1	(0.10,1,0.13,0.13,0.10,0.11,0.10,0.12)	1.05	53.20	1.95	0.69
	200	0	(0.06,1,0.17,0.29,0.16,0.16,0.14,0.14)	1.01	9.04	1.01	0.82
		0.25	(0.52,1,0.12,0.39,0.22,0.15,0.15,0.12)	0.97	9.98	1.02	0.86
		0.5	(0.55,1,0.13,0.44,0.25,0.11,0.14,0.15)	0.95	15.09	1.05	0.88
		0.75	(0.52,1,0.11,0.35,0.28,0.12,0.16,0.11)	0.95	26.90	1.16	0.87
		1	(0.09,1,0.10,0.13,0.09,0.13,0.13,0.13)	1.01	57.73	1.61	0.84
	500	0	(0.03,1,0.26,0.51,0.14,0.11,0.13,0.09)	1.01	8.65	1.02	0.94
		0.25	(0.63,1,0.07,0.64,0.31,0.10,0.12,0.07)	0.98	10.96	1.01	0.97
		0.5	(0.70,1,0.08,0.73,0.34,0.11,0.11,0.07)	0.96	16.04	1.01	0.98
		0.75	(0.61,1,0.08,0.58,0.35,0.19,0.12,0.09)	0.97	23.66	1.07	0.98
		1	(0.08,1,0.10,0.13,0.08,0.19,0.16,0.14)	1.01	54.30	1.29	0.96
0.4	100	0	(0.11,0.33,0.22,0.30,0.23,0.24,0.20,0.18)	1.02	1.46	1.05	0.65
		0.25	(0.35,0.38,0.18,0.38,0.25,0.17,0.17,0.19)	1.01	1.31	1.12	0.68
		0.5	(0.39,0.38,0.16,0.33,0.27,0.13,0.18,0.21)	1.01	1.48	1.22	0.70
		0.75	(0.37,0.48,0.15,0.29,0.20,0.11,0.18,0.22)	0.94	1.81	1.34	0.70
		1	(0.09,0.72,0.15,0.16,0.11,0.11,0.09,0.16)	1.05	4.16	1.95	0.67
	200	0	(0.09,0.42,0.16,0.35,0.20,0.18,0.14,0.15)	1.00	1.44	1.01	0.81
		0.25	(0.41,0.49,0.14,0.42,0.27,0.15,0.14,0.15)	0.98	1.40	1.02	0.84
		0.5	(0.37,0.46,0.13,0.45,0.28,0.12,0.15,0.18)	0.97	1.63	1.05	0.86
		0.75	(0.35,0.57,0.11,0.30,0.26,0.10,0.18,0.20)	0.96	2.15	1.16	0.85
		1	(0.08,0.94,0.10,0.14,0.09,0.12,0.12,0.14)	1.02	4.52	1.61	0.84
	500	0	(0.07,0.47,0.17,0.47,0.14,0.10,0.11,0.08)	1.01	1.43	1.02	0.94
		0.25	(0.38,0.60,0.08,0.57,0.31,0.12,0.09,0.10)	1.01	1.48	1.01	0.96
		0.5	(0.35,0.45,0.09,0.61,0.30,0.11,0.10,0.12)	0.97	1.64	1.01	0.96
		0.75	(0.42,0.72,0.09,0.44,0.30,0.18,0.13,0.21)	0.99	1.97	1.07	0.97
		1	(0.08,1,0.10,0.14,0.08,0.19,0.16,0.14)	1.01	4.22	1.29	0.96

Continued...

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}

Table B.33: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear31 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.15,0.96,0.24,1,0.21,0.21,0.18,0.98)	1.04	86.92	1.05	0.75
		0.25	(0.39,0.99,0.17,1,0.20,0.18,0.15,0.99)	1.00	112.13	1.06	0.78
		0.5	(0.44,0.99,0.14,1,0.20,0.14,0.16,1)	0.99	152.46	1.17	0.80
		0.75	(0.47,1,0.14,1,0.17,0.13,0.16,1)	0.99	240.40	1.28	0.81
		1	(0.07,1,0.10,1,0.08,0.09,0.10,1)	1.06	529.68	1.62	0.76
	200	0	(0.09,1,0.20,1,0.16,0.19,0.16,1)	1.00	100.46	1.01	0.86
		0.25	(0.50,1,0.15,1,0.22,0.17,0.15,1)	0.98	115.81	1.01	0.89
		0.5	(0.54,1,0.13,1,0.22,0.13,0.16,1)	0.97	181.41	1.05	0.91
		0.75	(0.52,1,0.12,1,0.25,0.15,0.17,1)	0.96	319.65	1.11	0.91
		1	(0.05,1,0.09,1,0.08,0.10,0.15,1)	1.01	602.37	1.41	0.87
	500	0	(0.04,1,0.30,1,0.13,0.14,0.16,1)	1.01	94.85	1.02	0.95
		0.25	(0.61,1,0.11,1,0.29,0.15,0.12,1)	0.98	126.10	1.00	0.97
		0.5	(0.66,1,0.10,1,0.33,0.15,0.12,1)	0.97	194.01	1.01	0.98
		0.75	(0.66,1,0.09,1,0.36,0.21,0.13,1)	0.98	287.11	1.04	0.98
		1	(0.07,1,0.07,1,0.07,0.17,0.21,1)	1.02	602.05	1.20	0.97
0.4	100	0	(0.12,0.29,0.16,0.73,0.19,0.17,0.16,0.32)	1.05	6.46	1.05	0.66
		0.25	(0.24,0.33,0.15,0.81,0.20,0.14,0.14,0.41)	1.01	7.16	1.06	0.70
		0.5	(0.26,0.34,0.13,0.80,0.22,0.13,0.18,0.51)	1.01	9.55	1.17	0.72
		0.75	(0.34,0.46,0.13,0.82,0.19,0.13,0.16,0.62)	0.99	14.67	1.28	0.75
		1	(0.09,0.64,0.11,0.84,0.13,0.09,0.10,0.62)	1.01	33.66	1.62	0.73
	200	0	(0.11,0.42,0.16,0.92,0.17,0.17,0.13,0.47)	1.01	7.07	1.01	0.83
		0.25	(0.32,0.47,0.13,0.93,0.22,0.17,0.13,0.56)	1.00	7.68	1.01	0.85
		0.5	(0.35,0.47,0.12,0.94,0.24,0.13,0.15,0.70)	0.98	11.50	1.05	0.88
		0.75	(0.41,0.62,0.11,0.94,0.24,0.11,0.17,0.80)	0.97	19.40	1.11	0.89
		1	(0.06,0.84,0.10,0.92,0.09,0.08,0.11,0.81)	1.02	38.47	1.41	0.87
	500	0	(0.06,0.49,0.22,1,0.15,0.13,0.13,0.78)	1.01	6.77	1.02	0.95
		0.25	(0.37,0.60,0.10,1,0.28,0.17,0.11,0.87)	1.00	8.34	1.00	0.96
		0.5	(0.38,0.52,0.10,1,0.29,0.20,0.11,0.96)	0.97	12.08	1.01	0.97
		0.75	(0.51,0.79,0.09,0.99,0.33,0.25,0.14,0.98)	0.99	17.39	1.04	0.98
		1	(0.07,0.99,0.08,0.99,0.07,0.17,0.20,0.97)	1.02	38.23	1.20	0.97

Table B.34: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear32 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.54,0.96,0.21,1,1,0.21,0.18,0.99)	1.04	286.82	1.03	0.82
		0.25	(0.01,1,0.15,1,1,0.15,0.15,1)	0.89	380.08	1.05	0.78
		0.5	(0.02,1,0.13,1,1,0.12,0.16,1)	0.82	501.89	1.12	0.79
		0.75	(0.01,1,0.11,1,1,0.11,0.14,1)	0.81	801.84	1.22	0.81
		1	(0.89,1,0.1,1,1,0.1,0.11,1)	1.03	1498.07	1.33	0.85
	200	0	(0.64,1,0.18,1,1,0.21,0.15,1)	1.01	344.37	1.02	0.90
		0.25	(0,1,0.13,1,1,0.14,0.14,1)	0.86	404.14	1.02	0.88
		0.5	(0,1,0.13,1,1,0.12,0.14,1)	0.81	622.08	1.05	0.88
		0.75	(0,1,0.13,1,1,0.08,0.18,1)	0.83	1092.25	1.10	0.90
		1	(0.99,1,0.1,1,1,0.1,0.14,1)	1.00	1771.08	1.21	0.93
	500	0	(0.85,1,0.28,1,1,0.14,0.14,1)	1.01	320.40	1.01	0.97
		0.25	(0,1,0.09,1,1,0.1,0.14,1)	0.85	438.38	1.01	0.95
		0.5	(0,1,0.14,1,1,0.1,0.17,1)	0.82	681.24	1.03	0.96
		0.75	(0,1,0.19,1,1,0.1,0.25,1)	0.80	1021.59	1.06	0.97
		1	(1,1,0.07,1,1,0.16,0.21,1)	1.01	1897.22	1.11	0.98

Continued...

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.4	100	0	(0.26,0.33,0.17,0.72,0.84,0.21,0.16,0.34)	1.05	19.03	1.03	0.70
		0.25	(0.11,0.33,0.16,0.81,0.93,0.15,0.16,0.44)	1.00	23.26	1.05	0.72
		0.5	(0.14,0.34,0.13,0.81,0.96,0.14,0.19,0.53)	0.95	30.63	1.12	0.75
		0.75	(0.21,0.47,0.13,0.82,0.98,0.12,0.15,0.66)	0.93	48.75	1.22	0.78
		1	(0.28,0.63,0.12,0.82,0.98,0.1,0.1,0.63)	0.96	93.96	1.33	0.78
	200	0	(0.3,0.4,0.16,0.93,0.97,0.2,0.13,0.5)	1.01	22.18	1.02	0.86
		0.25	(0.11,0.44,0.15,0.93,1,0.18,0.15,0.59)	0.98	25.26	1.02	0.86
		0.5	(0.14,0.41,0.13,0.94,1,0.14,0.14,0.75)	0.94	38.34	1.05	0.88
		0.75	(0.19,0.62,0.13,0.94,1,0.14,0.16,0.82)	0.94	66.43	1.10	0.89
		1	(0.33,0.85,0.11,0.92,1,0.09,0.12,0.83)	0.95	111.39	1.21	0.90
	500	0	(0.3,0.52,0.25,1,1,0.12,0.13,0.82)	1.00	20.79	1.01	0.95
		0.25	(0.05,0.56,0.15,1,1,0.23,0.12,0.89)	0.98	27.45	1.01	0.95
		0.5	(0.04,0.38,0.12,1,1,0.27,0.11,0.98)	0.95	41.61	1.03	0.96
		0.75	(0.1,0.82,0.1,0.99,1,0.42,0.14,0.99)	0.96	61.93	1.06	0.98
		1	(0.51,0.99,0.07,0.98,1,0.12,0.18,0.97)	0.98	118.89	1.11	0.98

Table B.35: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear33 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.91,0.93,0.85,0.99,0.99,0.99,0.97,0.98)	1.05	700.26	1.00	0.98
		0.25	(0.22,1,0.97,1,1,1,1,1)	0.93	910.19	1.00	0.92
		0.5	(0.06,1,1,1,1,1,1,1)	0.84	1135.78	1.00	0.93
		0.75	(0.04,1,0.99,1,1,1,1,1)	0.84	1656.57	1.00	0.94
		1	(0.89,1,1,1,1,1,1,1)	1.00	2826.16	1.00	0.99
	200	0	(0.98,1,0.98,1,1,1,1,1)	1.02	846.10	1.00	1.00
		0.25	(0.18,1,1,1,1,1,1,1)	0.89	999.63	1.00	0.96
		0.5	(0.04,1,1,1,1,1,1,1)	0.82	1491.56	1.00	0.95
		0.75	(0,1,1,1,1,1,1,1)	0.82	2501.75	1.00	0.96
		1	(0.98,1,1,1,1,1,1,1)	1.00	3651.46	1.00	1.00
	500	0	(1,1,1,1,1,1,1,1)	1.00	796.62	1.00	1.00
		0.25	(0.15,1,1,1,1,1,1,1)	0.88	1089.09	1.00	0.98
		0.5	(0,1,1,1,1,1,1,1)	0.81	1661.60	1.00	0.98
		0.75	(0,1,1,1,1,1,1,1)	0.78	2434.27	1.00	0.99
		1	(1,1,1,1,1,1,1,1)	1.00	4298.10	1.00	1.00
0.4	100	0	(0.61,0.35,0.26,0.75,0.84,0.65,0.39,0.38)	1.05	44.92	1.00	0.77
		0.25	(0.08,0.40,0.35,0.82,0.94,0.73,0.46,0.49)	1.00	55.66	1.00	0.77
		0.5	(0.10,0.37,0.38,0.82,0.97,0.79,0.55,0.60)	0.92	69.51	1.00	0.81
		0.75	(0.15,0.52,0.45,0.83,0.99,0.86,0.62,0.69)	0.87	101.32	1.00	0.85
		1	(0.34,0.69,0.55,0.82,0.97,0.94,0.49,0.68)	0.87	176.75	1.00	0.87
	200	0	(0.74,0.43,0.27,0.93,0.97,0.86,0.47,0.54)	1.01	53.33	1.00	0.90
		0.25	(0.07,0.50,0.40,0.92,1,0.92,0.62,0.64)	0.98	61.90	1.00	0.90
		0.5	(0.10,0.46,0.48,0.95,1,0.95,0.72,0.79)	0.94	91.85	1.00	0.92
		0.75	(0.14,0.65,0.63,0.94,1,0.98,0.80,0.86)	0.91	153.21	1.00	0.95
		1	(0.39,0.88,0.72,0.89,1,1,0.63,0.86)	0.91	228.76	1.00	0.96
	500	0	(0.91,0.51,0.27,1,1,1,0.69,0.84)	1.01	50.43	1.00	0.98
		0.25	(0.02,0.63,0.60,1,1,0.99,0.79,0.92)	0.98	67.56	1.00	0.97
		0.5	(0.02,0.45,0.75,1,1,1,0.91,0.98)	0.94	101.78	1.00	0.98
		0.75	(0.05,0.82,0.90,1,1,1,0.97,0.99)	0.92	148.73	1.00	0.99
		1	(0.57,0.99,0.95,0.98,1,1,0.92,0.99)	0.98	268.60	1.00	1.00

Table B.36: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from collinear34 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.67,0.15,0.10,0.08)	1.01	1.00	1.06	0.63
		0.25	(0.75,0.11,0.10,0.05)	1.02	1.00	1.14	0.64
		0.5	(0.77,0.15,0.06,0.02)	1.02	1.00	1.19	0.62
		0.75	(0.80,0.14,0.05,0.02)	1.01	1.00	1.28	0.62
		1	(0.88,0.10,0.02,0)	1.00	1.00	1.99	0.64
	200	0	(0.72,0.15,0.06,0.07)	1.01	1.00	1.03	0.80
		0.25	(0.74,0.13,0.08,0.05)	1.02	1.00	1.09	0.81
		0.5	(0.75,0.16,0.06,0.03)	1.01	1.00	1.08	0.81
		0.75	(0.80,0.13,0.06,0.01)	1.00	1.00	1.21	0.81
		1	(0.88,0.09,0.02,0)	1.01	1.00	1.51	0.81
	500	0	(0.70,0.16,0.06,0.08)	1.01	1.00	1.03	0.95
		0.25	(0.75,0.14,0.05,0.06)	1.01	1.00	1.02	0.95
		0.5	(0.68,0.18,0.07,0.06)	1.01	1.00	1.04	0.95
		0.75	(0.69,0.18,0.09,0.04)	1.01	1.00	1.09	0.95
		1	(0.85,0.08,0.06,0.01)	1.00	1.00	1.21	0.95
0.4	100	0	(0.67,0.15,0.10,0.08)	1.01	1.00	1.06	0.63
		0.25	(0.75,0.11,0.10,0.05)	1.02	1.00	1.14	0.64
		0.5	(0.77,0.15,0.06,0.02)	1.02	1.00	1.19	0.62
		0.75	(0.80,0.14,0.05,0.02)	1.01	1.00	1.28	0.62
		1	(0.88,0.10,0.02,0)	1.00	1.00	1.99	0.64
	200	0	(0.72,0.15,0.06,0.07)	1.01	1.00	1.03	0.80
		0.25	(0.74,0.13,0.08,0.05)	1.02	1.00	1.09	0.81
		0.5	(0.75,0.16,0.06,0.03)	1.01	1.00	1.08	0.81
		0.75	(0.80,0.13,0.06,0.01)	1.00	1.00	1.21	0.81
		1	(0.88,0.09,0.02,0)	1.01	1.00	1.51	0.81
	500	0	(0.70,0.16,0.06,0.08)	1.01	1.00	1.03	0.95
		0.25	(0.75,0.14,0.05,0.06)	1.01	1.00	1.02	0.95
		0.5	(0.68,0.18,0.07,0.06)	1.01	1.00	1.04	0.95
		0.75	(0.69,0.18,0.09,0.04)	1.01	1.00	1.09	0.95
		1	(0.85,0.08,0.06,0.01)	1.00	1.00	1.21	0.95

Table B.37: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from post-stratification1 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0,0,0.85,0.15)	1.00	32.05	1.02	0.63
		0.25	(0,0,0.86,0.14)	1.01	40.60	1.09	0.64
		0.5	(0,0,0.90,0.10)	1.02	60.74	1.10	0.62
		0.75	(0,0,0.93,0.07)	1.02	81.60	1.18	0.62
		1	(0,0,0.99,0.01)	1.00	164.05	1.49	0.64
	200	0	(0,0,0.83,0.17)	1.01	36.55	1.03	0.80
		0.25	(0,0,0.87,0.13)	1.01	43.77	1.05	0.81
		0.5	(0,0,0.91,0.09)	1.01	60.37	1.04	0.81
		0.75	(0,0,0.94,0.06)	1.00	87.59	1.14	0.81
		1	(0,0,0.98,0.02)	1.00	179.52	1.30	0.81
	500	0	(0,0,0.81,0.19)	1.01	37.47	1.02	0.95
		0.25	(0,0,0.83,0.17)	1.01	47.21	1.02	0.95
		0.5	(0,0,0.84,0.16)	1.01	54.45	1.03	0.95
		0.75	(0,0,0.88,0.12)	1.00	77.55	1.06	0.95
		1	(0,0,0.94,0.06)	1.00	187.23	1.14	0.95
0.4	100	0	(0,0,0.84,0.15)	1.00	2.77	1.02	0.63
		0.25	(0,0,0.86,0.14)	1.01	3.34	1.09	0.64
		0.5	(0,0,0.90,0.10)	1.02	4.39	1.10	0.62
		0.75	(0,0,0.93,0.07)	1.02	5.85	1.18	0.62
		1	(0,0,0.99,0.01)	1.00	11.04	1.49	0.64
	200	0	(0,0,0.83,0.17)	1.01	3.09	1.03	0.80
		0.25	(0,0,0.87,0.13)	1.01	3.64	1.05	0.81
		0.5	(0,0,0.91,0.09)	1.01	4.63	1.04	0.81
		0.75	(0,0,0.94,0.06)	1.00	6.53	1.14	0.81
		1	(0,0,0.98,0.02)	1.00	11.93	1.30	0.81
	500	0	(0,0,0.81,0.19)	1.01	3.24	1.02	0.95
		0.25	(0,0,0.83,0.17)	1.01	3.84	1.02	0.95

Continued...

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
		0.5	(0,0,0.84,0.16)	1.01	4.16	1.03	0.95
		0.75	(0,0,0.88,0.12)	1.00	5.59	1.06	0.95
		1	(0,0,0.94,0.06)	1.00	12.55	1.14	0.95

Table B.38: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from post-stratification2 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0,0,0,1)	1.00	131.82	1.00	0.63
		0.25	(0,0,0,1)	1.00	150.54	1.00	0.64
		0.5	(0,0,0,1)	1.00	234.37	1.00	0.62
		0.75	(0,0,0,1)	1.00	282.58	1.00	0.62
		1	(0,0,0,1)	1.00	464.55	1.00	0.64
	200	0	(0,0,0,1)	1.00	146.22	1.00	0.80
		0.25	(0,0,0,1)	1.00	170.54	1.00	0.81
		0.5	(0,0,0,1)	1.00	244.96	1.00	0.81
		0.75	(0,0,0,1)	1.00	320.72	1.00	0.81
		1	(0,0,0,1)	1.00	571.83	1.00	0.81
	500	0	(0,0,0,1)	1.00	151.81	1.00	0.95
		0.25	(0,0,0,1)	1.00	192.65	1.00	0.95
		0.5	(0,0,0,1)	1.00	220.46	1.00	0.95
		0.75	(0,0,0,1)	1.00	307.48	1.00	0.95
		1	(0,0,0,1)	1.00	702.22	1.00	0.95
0.4	100	0	(0,0,0,1)	1.00	8.86	1.00	0.63
		0.25	(0,0,0,1)	1.00	10.10	1.00	0.64
		0.5	(0,0,0,1)	1.00	14.90	1.00	0.62
		0.75	(0,0,0,1)	1.00	18.29	1.00	0.62
		1	(0,0,0,1)	1.00	29.56	1.00	0.64
	200	0	(0,0,0,1)	1.00	9.77	1.00	0.80
		0.25	(0,0,0,1)	1.00	11.52	1.00	0.81
		0.5	(0,0,0,1)	1.00	16.10	1.00	0.81
		0.75	(0,0,0,1)	1.00	21.20	1.00	0.81
		1	(0,0,0,1)	1.00	36.10	1.00	0.81
	500	0	(0,0,0,1)	1.00	10.38	1.00	0.95
		0.25	(0,0,0,1)	1.00	12.89	1.00	0.95
		0.5	(0,0,0,1)	1.00	14.43	1.00	0.95
		0.75	(0,0,0,1)	1.00	19.92	1.00	0.95
		1	(0,0,0,1)	1.00	44.72	1.00	0.95

Table B.39: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to true estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from post-stratification3 population of size $N = 1000$.

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}_{\mathrm{opt}}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.17,0.18,0.18)	(0,0,0)	1.05	1.00	1.07	0.86
		0.25	(0.18,0.21,0.19)	(0,0,0)	1.02	1.00	1.04	0.86
		0.5	(0.18,0.23,0.16)	(0,0,0)	1.04	1.00	1.04	0.85
		0.75	(0.18,0.22,0.16)	(0,0,0)	1.03	1.00	1.05	0.86
		1	(0.17,0.22,0.17)	(0,0,0)	1.02	1.00	1.03	0.85
	200	0	(0.20,0.17,0.15)	(0,0,0)	1.02	1.00	1.03	0.93
		0.25	(0.19,0.21,0.18)	(0,0,0)	1.02	1.00	1.03	0.93
		0.5	(0.18,0.24,0.19)	(0,0,0)	1.01	1.00	1.02	0.93
		0.75	(0.17,0.24,0.17)	(0,0,0)	1.02	1.00	1.03	0.93
		1	(0.17,0.29,0.16)	(0,0,0)	1.03	1.00	1.04	0.93
	500	0	(0.16,0.11,0.12)	(0,0,0)	1.01	1.00	1.01	0.98

Continued. . .

σ	n	r	$\mathrm{E}\,(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}_{\mathrm{opt}}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
		0.25	(0.19,0.23,0.20)	(0,1,1)	1.01	1.01	1.00	0.98
		0.5	(0.19,0.30,0.22)	(0,0,1)	1.01	1.00	1.00	0.98
		0.75	(0.25,0.29,0.21)	(0,1,0)	1.00	1.00	1.00	0.98
		1	(0.23,0.43,0.16)	(0,1,0)	1.00	1.00	1.00	0.98
0.4	100	0	(0.18,0.18,0.18)	(0,0,0)	1.04	1.00	1.05	0.85
		0.25	(0.17,0.17,0.18)	(0,0,0)	1.02	1.00	1.05	0.85
		0.5	(0.17,0.19,0.16)	(0,0,0)	1.04	1.00	1.05	0.85
		0.75	(0.14,0.18,0.13)	(0,0,0)	1.02	1.00	1.08	0.84
		1	(0.12,0.19,0.10)	(0,0,0)	1.01	1.00	1.12	0.84
	200	0	(0.18,0.16,0.16)	(0,0,0)	1.01	1.00	1.03	0.92
		0.25	(0.15,0.16,0.19)	(0,0,0)	1.01	1.00	1.03	0.92
		0.5	(0.14,0.17,0.18)	(0,0,0)	1.01	1.00	1.04	0.93
		0.75	(0.10,0.15,0.17)	(0,0,0)	1.01	1.00	1.05	0.92
		1	(0.11,0.21,0.11)	(0,0,0)	1.01	1.00	1.12	0.92
	500	0	(0.15,0.19,0.09)	(0,0,0)	1.00	1.00	1.00	0.98
		0.25	(0.14,0.13,0.26)	(0,0,1)	1.01	1.00	1.01	0.98
		0.5	(0.12,0.14,0.25)	(1,0,1)	1.01	1.00	1.01	0.98
		0.75	(0.14,0.15,0.26)	(0,1,0)	1.00	1.00	1.01	0.98
		1	(0.16,0.30,0.17)	(0,0,0)	1.01	1.00	1.05	0.98

Table B.40: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to optimal estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from quadratic population of size $N = 1000$ based on linear regression.

σ	n	r	$\mathrm{E}\,(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}_{\mathrm{opt}}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}
0.1	100	0	(0.10,0.73,0.10,0.12,0.73,0.13,0.14,0.12,0.18)	(0,1,0,0,1,0,0,0,0)	1.52	2.50	1.08	0.66
		0.25	(0.12,0.74,0.13,0.12,0.74,0.12,0.15,0.14,0.19)	(0,1,0,0,1,0,0,0,0)	1.77	3.54	1.16	0.66
		0.5	(0.09,0.72,0.10,0.11,0.72,0.10,0.14,0.13,0.18)	(0,1,0,0,1,0,0,0,0)	1.99	4.25	1.20	0.66
		0.75	(0.10,0.75,0.08,0.10,0.75,0.10,0.13,0.10,0.19)	(0,1,0,0,1,0,0,0,0)	2.37	5.94	1.51	0.66
		1	(0.09,0.72,0.08,0.10,0.72,0.11,0.11,0.12,0.17)	(0,1,0,0,1,0,0,0,0)	4.30	12.13	2.08	0.66
	200	0	(0.08,0.92,0.09,0.10,0.92,0.10,0.16,0.14,0.18)	(0,1,0,0,1,0,0,0,0)	1.12	2.95	1.05	0.83
		0.25	(0.10,0.94,0.09,0.10,0.94,0.10,0.19,0.15,0.23)	(0,1,0,0,1,0,0,0,1)	1.24	3.48	1.08	0.84
		0.5	(0.08,0.92,0.10,0.10,0.92,0.10,0.16,0.12,0.21)	(0,1,0,0,1,0,0,0,0)	1.31	4.64	1.11	0.83
		0.75	(0.08,0.94,0.09,0.07,0.94,0.07,0.14,0.09,0.21)	(0,1,0,0,1,0,0,0,0)	1.50	6.81	1.25	0.83
		1	(0.06,0.92,0.08,0.07,0.92,0.08,0.11,0.07,0.19)	(0,1,0,0,1,0,0,0,0)	1.93	12.81	1.63	0.82
	500	0	(0.07,1,0.05,0.08,1,0.05,0.11,0.09,0.09)	(0,1,0,0,1,0,0,0,1)	1.00	2.68	1.02	0.96
		0.25	(0.06,1,0.09,0.07,1,0.11,0.15,0.12,0.24)	(0,1,0,0,1,0,0,1,1)	1.01	3.66	1.03	0.96
		0.5	(0.06,1,0.10,0.06,1,0.09,0.10,0.09,0.21)	(0,1,0,0,1,0,0,1,0)	1.01	4.34	1.06	0.96
		0.75	(0.04,1,0.11,0.06,1,0.07,0.09,0.08,0.20)	(0,1,0,0,1,0,0,0,0)	1.01	6.66	1.13	0.96
		1	(0.04,0.99,0.08,0.05,0.99,0.09,0.10,0.09,0.19)	(0,1,0,0,1,0,0,0,0)	1.06	12.62	1.26	0.96
0.4	100	0	(0.10,0.31,0.08,0.10,0.39,0.13,0.11,0.11,0.13)	(0,1,0,0,1,0,0,0,0)	1.06	1.02	1.08	0.62
		0.25	(0.10,0.31,0.10,0.10,0.36,0.10,0.12,0.12,0.14)	(0,1,0,0,1,0,0,0,0)	1.11	1.11	1.16	0.62
		0.5	(0.08,0.29,0.07,0.09,0.29,0.07,0.10,0.09,0.13)	(0,1,0,0,1,0,0,0,0)	1.15	1.14	1.20	0.61
		0.75	(0.07,0.33,0.07,0.07,0.34,0.06,0.11,0.10,0.10)	(0,1,0,0,1,0,0,0,0)	1.17	1.18	1.51	0.61
		1	(0.07,0.36,0.07,0.06,0.37,0.06,0.09,0.08,0.11)	(0,1,0,0,1,0,0,0,0)	1.27	1.43	2.08	0.61
	200	0	(0.08,0.49,0.09,0.10,0.52,0.09,0.10,0.14,0.11)	(0,1,0,0,1,0,0,0,0)	1.08	1.09	1.05	0.80
		0.25	(0.08,0.50,0.10,0.10,0.52,0.09,0.12,0.12,0.14)	(0,1,0,0,1,0,0,0,1)	1.08	1.07	1.08	0.80
		0.5	(0.09,0.44,0.09,0.09,0.44,0.08,0.11,0.10,0.15)	(0,1,0,0,1,0,0,0,0)	1.10	1.15	1.11	0.80
		0.75	(0.07,0.50,0.08,0.05,0.50,0.06,0.08,0.08,0.16)	(0,1,0,0,1,0,0,0,0)	1.14	1.30	1.25	0.81
		1	(0.06,0.48,0.07,0.06,0.48,0.07,0.10,0.06,0.14)	(0,1,0,0,1,0,0,0,0)	1.23	1.49	1.63	0.80
	500	0	(0.07,0.82,0.05,0.09,0.82,0.05,0.09,0.08,0.06)	(0,1,0,0,1,0,0,0,1)	1.00	1.06	1.02	0.95
		0.25	(0.06,0.78,0.13,0.08,0.78,0.10,0.12,0.13,0.20)	(0,1,0,0,1,0,0,1,1)	1.03	1.09	1.03	0.96
		0.5	(0.07,0.66,0.10,0.06,0.66,0.06,0.10,0.10,0.22)	(0,1,0,0,1,0,0,1,0)	1.05	1.13	1.06	0.95
		0.75	(0.05,0.76,0.12,0.07,0.76,0.08,0.09,0.08,0.22)	(0,1,0,0,1,0,0,0,0)	1.06	1.24	1.13	0.96
		1	(0.05,0.64,0.08,0.07,0.64,0.06,0.12,0.10,0.28)	(0,1,0,0,1,0,0,0,0)	1.18	1.51	1.26	0.95

Continued...

σ	n	r	$\mathrm{E}(\widehat{\boldsymbol{c}}_{\mathrm{CV}})$	$\boldsymbol{c}_{\mathrm{opt}}$	δ_{CV}	δ_{π}	δ_{reg}	R_{w}

Table B.41: Expected $\widehat{\boldsymbol{c}}_{\mathrm{CV}}$ and ratio of MSE of cross-validation (CV), π, and regression (reg) estimators to optimal estimator, and ratio of variance of weights of CV estimator to regression estimator from 1000 replications of stratified simple random sampling from quadratic population of size $N = 1000$ based on quadratic regression.

BIBLIOGRAPHY

Bibliography

Breidt, F. J., G. Claeskens, and J. D. Opsomer (2005). Model-assisted estimation for complex surveys using penalised splines. *Biometrika 92*(4), 831–846.

Breidt, F. J. and J. D. Opsomer (2000). Local polynomial regression estimators in survey sampling. *Annals of Statistics 28*, 1026–1053.

Brumback, B. A. and J. A. Rice (1998). Smoothing spline models for the analysis of nested and crossed samples of curves (c/r: P976-994). *Journal of the American Statistical Association 93*, 961–976.

Cochran, W. G. (1977). *Sampling Techniques.* New York: John Wiley & Sons.

Coull, B. A., D. Ruppert, and M. P. Wand (2001). Simple incorporation of interactions into additive models. *Biometrics 57*, 539–545.

Eilers, P. H. C. and B. D. Marx (1996). Flexible smoothing with B-splines and penalties. *Stat. Science 11*(2), 89–121.

Friedman, J. H. (1991). Multivariate adaptive regression splines (with discussion). *Ann. Statist. 19*, 1–141.

Friedman, J. H. and B. W. Silverman (1989). Flexible parsimonious smoothing and additive modeling (with discussion). *Technometrics 31*, 3–39.

Fuller, W. A. (1996). *Introduction to Statistical Time Series.* New York: John Wiley & Sons.

Green, D. J. and B. W. Silverman (1994). *Nonparametric Regression and Generalized Linear Models.* Chapman & Hall.

Horvitz, D. G. and D. J. Thompson (1952). A generalization of sampling without replacement from a finite universe. *Journal of the American Statistical Association 47*, 663–685.

Kauermann, G. (2005). A note on smoothing parameter selection for penalised spline smoothing. *Journal of Statistical Planing and Inference 127*, 53–69.

Opsomer, J. D., G. Claeskens, M. G. Ranalli, G. Kauermann, and F. J. Breidt (2008). Nonparametric small area estimation using penalized spline regression. *Journal of the Royal Statistical Society, Series B 70*, 265–286.

Opsomer, J. D. and C. P. Miller (2005). Selecting the amount of smoothing in nonparametric regression estimation for complex surveys. *Nonparametric Statistics 17*, 593–611.

Parise, H., D. Ruppert, L. Ryan, and M. P. Wand (2001). Incorporation of historical controls using semiparametric mixed models. *Applied Statistics 50*, 31–42.

Ruppert, D. (2002). Selecting the number of knots for penalized splines. *Journal of Computational and Graphical Statistics 11*, 735–757.

Ruppert, R., M. Wand, and R. Carroll (2003). *Semiparametric Regression*. Cambridge University Press.

Särndal, C. E. (1982). Implications of survey design for generalized regression estimation of linear functions. *Journal of Statistical Planning and Inference 7*, 155–170.

Särndal, C. E., B. Swensson, and J. Wretman (1989). The weighted residual technique for estimating the variance of the general regression estimator of the finite population total. *Biometrika 76*, 527–537.

Särndal, C. E., B. Swensson, and J. Wretman (1992). *Model Assisted Survey Sampling*. New York: Springer-Verlag.

Verbyla, A. P., B. R. Cullis, M. G. Kenward, and S. J. Welham (1999). The analysis of designed experiments and longitudinal data by using smoothing splines. *Applied Statistics 48*, 269–311.

Wand, M. (1999). On the optimal amount of smoothing in penalised spline regression. *Biometrika 86*, 936–940.

Wand, M. (2003). Smoothing and mixed models. *Computational Statistics 18*, 223–249.

Wand, M. P. (2000). A comparison of regression spline smoothing procedures. *Computational Statistics 15*, 443–462.

Zheng, H. and R. J. A. Little (2003). Penalized spline model-based estimation of finite population total from probability-proportional-to-size samples. *Journal of Official Statistics 19*, 99–117.